

## Stabilità dell'integrabilità Hamiltoniana: teorema KAM e serie di Lindstedt

Complemento di *Giovanni Gallavotti* al Cap.X di [1], v. 1.9

### 1 Introduzione

Moltissimi sono i problemi di Meccanica che trattano dei moti di sistemi Hamiltoniani la cui funzione di Hamilton è “vicina” a quella di un sistema analiticamente integrabile.

Un sistema a  $\ell$  gradi di libertà e con Hamiltoniana  $H(\vec{p}, \vec{q})$  analitica sullo spazio delle fasi è analiticamente integrabile su una regione  $W$  dello spazio delle fasi se

(1) i punti di  $W$  sono rappresentabili via  $\ell$  coordinate  $\vec{A} = (A_1, \dots, A_\ell) \in \mathbb{R}^\ell$ , dette *azioni*, e  $\ell$  angoli  $\vec{\alpha} = (\alpha_1, \dots, \alpha_\ell) \in [0, 2\pi]^\ell$  nel senso che i punti  $(\vec{p}, \vec{q}) \in W$  possono essere rappresentati, a mezzo di una trasformazione analitica<sup>1</sup> di coordinate  $\vec{p} = \vec{P}(\vec{A}, \vec{\alpha})$ ,  $\vec{q} = \vec{Q}(\vec{A}, \vec{\alpha})$  con  $(\vec{A}, \vec{\alpha}) \in U \times T^\ell$  ove  $U \subset \mathbb{R}^\ell$  è un dominio chiuso di un aperto connesso e limitato<sup>2</sup> e  $T^\ell$  è il toro  $\ell$ -dimensionale  $T^\ell$ ,

(2) la trasformazione  $(\vec{P}(\vec{A}, \vec{\alpha}), \vec{Q}(\vec{A}, \vec{\alpha})) \longleftrightarrow (\vec{A}, \vec{\alpha})$  è *completamente canonica*, [1],

(3) nelle suddette coordinate “azione-angolo” l'Hamiltoniana dipende dalle sole  $\vec{A}$ :  $H(\vec{P}(\vec{A}, \vec{\alpha}), \vec{Q}(\vec{A}, \vec{\alpha})) = h(\vec{A})$ .

Quindi, essendo la trasformazione in coordinate azione-angolo completamente canonica, le equazioni del moto con dato iniziale  $(\vec{A}_0, \vec{\alpha}_0)$ , denotando sempre con  $\vec{\partial}_{\vec{x}} \stackrel{def}{=} (\frac{\partial}{\partial x_1}, \dots, \frac{\partial}{\partial x_n})$  per  $\vec{x} \in \mathbb{R}^n$ , sono semplicemente  $\dot{\vec{A}} = \vec{0}$ ,  $\dot{\vec{\alpha}} =$

---

<sup>1</sup>Una funzione di  $\vec{x} \in U \subset \mathbb{R}^n$  definita su  $U$ , chiusa di un aperto connesso e limitato, e con tutte le derivate ben definite nei punti interni e continue fin sul bordo è analitica se è sviluppabile in serie di Taylor attorno ad ogni punto di  $U$  (e quindi è olomorfa in un aperto di  $C^n$  che contiene  $U$ ). Un sistema di coordinate su  $W$  è analiticamente regolare se punti corrispondenti in carte diverse hanno coordinate legate da funzioni analitiche.

<sup>2</sup>Senza perdita di generalità si potrà supporre che  $U$  sia una sfera in  $\mathbb{R}^\ell$  perché i problemi di stabilità che discuteremo hanno carattere locale.

$\vec{\partial}_{\vec{A}}h(\vec{A})$ , e (banalmente) risolte da

$$\vec{A}(t) = \vec{A}_0, \quad \vec{\alpha}(t) = \vec{\alpha}_0 + \vec{\omega}(\vec{A}_0)t, \quad \vec{\omega}(\vec{A}) \stackrel{\text{def}}{=} \vec{\partial}_{\vec{A}}h(\vec{A}) \quad (1.1)$$

ove  $\vec{\omega}(\vec{A})$  vien chiamato “*vettore di rotazione*” o “*spettro di frequenza*” dei moti con dato iniziale  $\vec{A}$ .

Si può dire che un sistema integrabile ammette *variabili azione-angolo*,  $\vec{A}, \vec{\alpha}$ , che evolvono nel tempo: le prime restando costanti e le seconde quali angoli che ruotano uniformemente con velocità angolari  $\omega_i(\vec{A})$  (frequenza  $2\pi\omega_i(\vec{A})^{-1}$ ):  $(\vec{A}(t), \vec{\alpha}(t)) = (\vec{A}, \vec{\alpha} + \vec{\omega}(\vec{A})t)$ . Quindi l’evoluzione di una qualsiasi grandezza  $F(\vec{A}, \vec{\alpha})$ , analitica su  $W$ , è una funzione  $t \rightarrow F(\vec{A}(t), \vec{\alpha}(t))$  *quasi periodica* del tempo.<sup>3</sup>

Si dice anche che un sistema Hamiltoniano a  $\ell$  gradi di libertà è integrabile su  $W$  se  $W$  può essere pensato come “*foliato*” in tori invarianti  $\ell$ -dimensionali percorsi da moti quasi periodici parametrizzati da  $\ell$  azioni.

Se  $H_\varepsilon$  è analitica e “vicina” ad  $H$  allora, calcolandola nelle variabili azione-angolo che integrano  $H$ , assume la forma

$$H_\varepsilon(\vec{A}, \vec{\alpha}) = h(\vec{A}) + \varepsilon f(\vec{A}, \vec{\alpha}; \varepsilon) \quad (1.2)$$

con  $\varepsilon$  “piccolo” e  $f$  analitica su  $W = U \times T^\ell$  e in  $\varepsilon$  vicino a 0.

Quindi lo studio dei moti di sistemi prossimi a sistemi integrabili si riduce, senza perdita di generalità, al caso della Eq.(1.2) e alle equazioni:

$$\dot{\vec{A}} = -\varepsilon \vec{\partial}_{\vec{\alpha}}f(\vec{A}, \vec{\alpha}, \varepsilon), \quad \dot{\vec{\alpha}} = \vec{\omega}(\vec{A}) + \varepsilon \vec{\partial}_{\vec{A}}f(\vec{A}, \vec{\alpha}, \varepsilon) \quad (1.3)$$

Queste equazioni si possono studiare cercandone una soluzione in serie di potenze in  $\varepsilon$ . Ma Poincaré dimostrò che *genericamente* i moti governati da  $H_\varepsilon$  sono di natura diversa dai moti imperturbati (ossia non sono tutti quasi

---

<sup>3</sup>Una funzione  $\varphi(t)$  è quasi periodica in  $t$ , a  $\ell$  frequenze, se esistono  $\ell$  velocità angolari  $\vec{\omega} = (\omega_1, \dots, \omega_\ell)$  e  $\varphi(t) = \sum_{\vec{\nu} \in \mathbb{Z}^\ell} \varphi_{\vec{\nu}} e^{i\vec{\omega} \cdot \vec{\nu} t}$  con  $\sum_{\vec{\nu} \in \mathbb{Z}^\ell} |\varphi_{\vec{\nu}}| < \infty$ . Allora  $t \rightarrow F(\vec{A}(t), \vec{\alpha}(t))$  è quasi periodica perché ammette la rappresentazione  $F(\vec{A}(t), \vec{\alpha}(t)) = \sum_{\vec{\nu} \in \mathbb{Z}^\ell} F_{\vec{\nu}}(\vec{A}) e^{i\vec{\nu} \cdot (\vec{\alpha} + \vec{\omega}(\vec{A})t)}$  ove  $F_{\vec{\nu}}(\vec{A})$  sono i coefficienti di Fourier della funzione  $F(\vec{A}, \vec{\alpha})$  degli angoli  $\vec{\alpha}$  con  $|F_{\vec{\nu}}(\vec{A})| < C e^{-\kappa|\vec{\nu}|}$  per  $C, \kappa > 0$  opportuni. L’esistenza di  $C, \kappa$  consegue dal teorema di Cauchy per le funzioni olomorfe.

periodici)<sup>4</sup>: per  $\varepsilon$  piccolo, se  $\neq 0$ , non può esistere un cambiamento canonico di coordinate che, per  $\varepsilon$  abbastanza piccolo, sia:

(1) circa l'identità, cioè espresso, da funzioni  $\vec{\Xi}_\varepsilon, \vec{\Delta}_\varepsilon$  analitiche in  $\vec{A}', \vec{\alpha}', \varepsilon$ , come

$$\vec{A} = \vec{A}' + \varepsilon \vec{\Xi}_\varepsilon(\vec{A}', \vec{\alpha}'), \quad \vec{\alpha} = \vec{\alpha}' + \varepsilon \vec{\Delta}_\varepsilon(\vec{A}', \vec{\alpha}') \quad (1.4)$$

(2) a valori in  $U \times T^\ell$  e definito per  $|\varepsilon|$  piccolo e  $(\vec{A}', \vec{\alpha}') \in U' \times T^\ell$ , con  $U' \subset U$  e contenente tutti i punti a distanza dalla frontiera di  $U$  maggiore o uguale a una quantità  $\lambda$  prefissata,<sup>5</sup>

(3) e tale che sia

$$H_\varepsilon(\vec{A}' + \varepsilon \vec{\Xi}_\varepsilon(\vec{A}', \vec{\alpha}'; \varepsilon), \vec{\alpha}' + \varepsilon \vec{\Delta}_\varepsilon(\vec{A}', \vec{\alpha}'; \varepsilon)) = h_\varepsilon(\vec{A}') \quad (1.5)$$

con una funzione  $h_\varepsilon(\vec{A}')$ , analitica per  $\varepsilon$  piccolo e per  $\vec{A}' \in U'$ .

In altre parole non è possibile che una perturbazione *generica* mantenga il sistema perturbato ancora integrabile. La dimostrazione è particolarmente semplice, [1], se la matrice Hessiana  $\partial_{A_i A_j}^2 h(\vec{A})$  è non degenere ( $\det \vec{\partial}^2 h(\vec{A}) \neq 0$ ): si veda Appendice A.

L'equazione Eq.(1.5) per  $h_\varepsilon, \vec{\Xi}_\varepsilon, \vec{\Delta}_\varepsilon$  con la condizione che la Eq.(1.4) sia una trasformazione canonica è soddisfatta se esiste una funzione "generatrice"  $\Phi_\varepsilon(\vec{A}, \vec{\alpha})$  di  $\varepsilon, \vec{A}, \vec{\alpha}$  analitica per  $\varepsilon$  piccolo e per  $(\vec{A}, \vec{\alpha}) \in U' \times T^\ell$  tale che la Eq.(1.5) sia implicata da

$$\vec{A} = \vec{A}' + \varepsilon \vec{\partial}_{\vec{\alpha}} \Phi_\varepsilon(\vec{A}', \vec{\alpha}), \quad \vec{\alpha}' = \vec{\alpha} + \varepsilon \vec{\partial}_{\vec{A}'} \Phi_\varepsilon(\vec{A}', \vec{\alpha}) \quad (1.6)$$

cosicché la Eq.(1.5) diviene:

$$H_\varepsilon(\vec{A}' + \varepsilon \vec{\partial}_{\vec{\alpha}} \Phi_\varepsilon(\vec{A}', \vec{\alpha}), \vec{\alpha}) = h_\varepsilon(\vec{A}') \quad (1.7)$$

che è detta *equazione di Hamilton-Jacobi* per l'integrazione perturbativa: perché l'esistenza di soluzione alla Eq.(1.7) definisce implicitamente  $\vec{\Xi}_\varepsilon, \vec{\Delta}_\varepsilon$

---

<sup>4</sup>Senza riferirci alla definizione "formale" di proprietà genericamente vera si può qui interpretare la genericità dicendo che la proprietà in questione vale per le funzioni perturbatrici  $f$  che hanno tutti i coefficienti di Fourier  $f_{\vec{\nu}}(\vec{A})$  non nulli per  $\vec{\nu} \neq \vec{0}$  nell'intorno di un punto di  $U$ .

<sup>5</sup>Qui  $\lambda$  è necessario perché se  $\varepsilon \neq 0$  il moto dei punti troppo vicini alla frontiera di  $U \times T^\ell$  potrebbe comunque arrivare alla frontiera in un tempo finito pur restando il sistema integrabile abbastanza lontano dalla frontiera.

che soddisfano la Eq.(1.5) e quindi implica l'integrabilità analitica su  $U' \times \ell$  dell'Hamiltoniana  $H_\varepsilon$  per  $\varepsilon$  piccolo.

Per il risultato citato di Poincaré la Eq.(1.7) *non è in generale risolubile* “globalmente” su domini  $U' \times T^\ell$  con  $U' \simeq U$ . Il problema che si pone, visto questo risultato negativo, è cosa sopravviva della foliazione in tori invarianti dello spazio delle fasi e dei moti quasi periodici che si svolgono su di essi se i moti sono generati dall'Hamiltoniana  $H_\varepsilon$  in Eq.(1.2) con  $\varepsilon \neq 0$ .

## 2 Teoria perturbativa classica

La teoria perturbativa classica risale al XVIII secolo e nel XIX secolo fu riformulata a mezzo del formalismo Hamiltoniano che consente di darne una elegante formulazione algoritmica di applicabilità generale (e quasi meccanica). La illustriamo qui in dettaglio perché la sua piena comprensione è illuminante per accedere agli sviluppi “rivoluzionari”, [2, 3, 4, 5], avvenuti nella seconda metà del XX secolo.

Se il dato iniziale è  $(\vec{A}_0, \vec{\alpha}_0)$  e se si vuole *solo* calcolare il moto *al variare del tempo in un intervallo*  $[0, T]$  *arbitrario ma prefissato* allora si potrebbe pensare di procedere così. Immaginiamo di sviluppare in serie di  $\varepsilon$  la soluzione dell'equazione del moto, il che è lecito per la dipendenza analitica da dati iniziali e parametri delle equazioni differenziali ordinarie.

Dai teoremi sulla dipendenza continua dai dati iniziali si deduce che, entro il tempo prefissato  $T$ , il moto è ben approssimato *a meno di*  $O(\varepsilon)$  dal moto quasi periodico imperturbato: ossia con  $\varepsilon = 0$  e  $\vec{A}(t) = \vec{A}_0$ ,  $\vec{\alpha}(t) = \vec{\alpha}_0 + \vec{\omega}_0 t$ ; allora integrando la Eq.(1.3) e trascurando quantità di  $O(\varepsilon^2)$ , si troverà, se  $\vec{\omega}_0 \stackrel{def}{=} \vec{\omega}(\vec{A}_0)$ :

$$\begin{aligned} \vec{A}(t) &= \vec{A}_0 - \varepsilon \int_0^t \vec{\partial}_{\vec{\alpha}} f(\vec{A}_0, \vec{\alpha}_0 + \vec{\omega}_0 \tau) d\tau + O(\varepsilon^2) \\ \vec{\alpha}(t) &= \vec{\alpha}_0 + \vec{\omega}_0 t + \int_0^t \left[ ((\vec{A}(\tau) - \vec{A}_0) \cdot \vec{\partial}_{\vec{A}}) \vec{\omega}(\vec{A}_0) \right. \\ &\quad \left. + \varepsilon \vec{\partial}_{\vec{A}} f(\vec{A}_0, \vec{\alpha}_0 + \vec{\omega}_0 \tau) \right] d\tau + O(\varepsilon^2) \end{aligned} \quad (2.1)$$

e qui, come in seguito se non è indicato esplicitamente, l'argomento  $\varepsilon$  di  $f$  è  $\varepsilon = 0$  e viene sottinteso.

Come pervenire ad approssimazioni migliori ossia on errore  $< O(\varepsilon^2)$ ? La questione è di grande importanza perché attorno ad essa volgono i calcoli più semplici e classici della Meccanica Celeste. Per procedere conviene prima

esaminare in dettaglio la struttura dell'approssimazione in Eq.(2.1) a  $O(\varepsilon)$  che, a tale ordine, può essere calcolata *esattamente e elementarmente*.

Sia  $f(\vec{A}, \vec{\alpha}, \varepsilon) = \sum_{\vec{\nu} \in Z^\ell} f_{\vec{\nu}}(\vec{A}, \varepsilon) e^{i\vec{\nu} \cdot \vec{\alpha}}$ , ove  $\vec{\nu} \in Z^\ell$ ,  $\vec{\nu} = (\nu_1, \dots, \nu_\ell)$  sono vettori a componenti intere, la rappresentazione di Fourier della  $f$ : allora eseguendo esplicitamente (termine a termine) gli integrali nella Eq.(2.1) questa diventa, *a meno di un errore*  $O(\varepsilon^2)$  (sempre entro il tempo prefissato  $T$ ):

$$\begin{aligned} \vec{A}(t) - \vec{A}_0 &= -\varepsilon \sum_{\vec{\nu} \neq \vec{0}} \frac{i\vec{\nu} f_{\vec{\nu}}(\vec{A}_0)}{i\vec{\omega}_0 \cdot \vec{\nu}} e^{i\vec{\nu} \cdot \vec{\alpha}_0} (e^{i\vec{\omega}_0 \cdot \vec{\nu} t} - 1) \\ \vec{\alpha}(t) - \vec{\alpha}_0 &= \left( \vec{\omega}_0 + \varepsilon \vec{\partial}_{\vec{A}} f_{\vec{0}}(\vec{A}_0) + \varepsilon \sum_{\vec{\nu} \neq \vec{0}} \frac{f_{\vec{\nu}}(\vec{A}_0) i(\vec{\nu} \cdot \vec{\partial}_{\vec{A}}) \vec{\omega}(\vec{A}_0)}{i\vec{\omega}_0 \cdot \vec{\nu}} e^{i\vec{\nu} \cdot \vec{\alpha}_0} \right) t \quad (2.2) \\ &+ \varepsilon \sum_{\vec{\nu} \neq \vec{0}} e^{i\vec{\nu} \cdot \vec{\alpha}_0} (e^{i\vec{\omega}_0 \cdot \vec{\nu} t} - 1) \left( \frac{\vec{\partial}_{\vec{A}} f_{\vec{\nu}}(\vec{A}_0)}{i\vec{\omega}_0 \cdot \vec{\nu}} - f_{\vec{\nu}}(\vec{A}_0) \frac{i(\vec{\nu} \cdot \vec{\partial}_{\vec{A}}) \vec{\omega}(\vec{A}_0)}{(i\vec{\omega}_0 \cdot \vec{\nu})^2} \right) \end{aligned}$$

ove  $f_{\vec{0}}(\vec{A})$  è la componente  $\vec{0}$  di Fourier e quindi è *la media sugli angoli*  $f_{\vec{0}}(\vec{A}) = \int_{T^\ell} \frac{d\vec{\alpha}}{(2\pi)^\ell} f(\vec{A}, \vec{\alpha})$  della funzione perturbatrice  $f$ .

Certo si deve verificare che le serie convergano: e qui si incontra la prima difficoltà addirittura già presente nel caso in cui la perturbazione  $f$  sia un polinomio trigonometrico e abbia solo un numero finito di componenti di Fourier e quindi le somme in Eq.(2.2) vertano su un numero finito di  $\vec{\nu}$ .

Invero si devono escludere le “risonanze” che sono i casi in cui le velocità angolari  $\vec{\omega}_0$  sono *risonanti* cioè sono razionalmente dipendenti nel senso che esiste un vettore  $\vec{\nu} \neq \vec{0}$  a componenti intere tale che  $\vec{\omega}_0 \cdot \vec{\nu} = 0$ : questo infatti potrebbe impedire di definire le serie in Eq.(2.2) a causa della divisione per 0 che ne conseguirebbe.

Anzi anche se  $\vec{\omega}_0 \cdot \vec{\nu} \neq 0$  per tutti i  $\vec{\nu} \neq \vec{0}$  tale quantità può essere talmente vicina a 0 e per talmente tanti  $\vec{\nu} \neq \vec{0}$  che le serie in Eq.(2.2), pur definite termine a termine, sono divergenti e quindi non definite. Inoltre se  $h(\vec{A})$  ha matrice Hessiana non singolare ( $\det \vec{\partial}_{\vec{A}}^2 h(\vec{A}) \neq 0$ ) e se  $f$  ha tutte le componenti di Fourier non nulle, nell'intorno di un qualsiasi  $\vec{\omega} = \vec{\partial}_{\vec{A}} h(\vec{A})$  esiste un insieme denso di vettori  $\vec{\omega}_0$  *risonanti*, ossia tali che per qualche  $\vec{\nu} \neq \vec{0}$  (dipendente da  $\vec{\omega}_0$ ) si ha  $\vec{\nu} \cdot \vec{\omega}_0 = 0$ , e per questi la Eq.(2.2) non ha senso!

Nelle applicazioni spesso la perturbazione è un polinomio trigonometrico ossia  $f_{\vec{\nu}}$  ha solo un numero finito di componenti di Fourier e quindi l'analisi sopra discussa ha senso ed è utile almeno se  $O(\varepsilon^2)$  è una stima di errore

buona per gli scopi, nel tempo  $T$  di interesse, della particolare applicazione e se nessuno dei denominatori coinvolti nelle serie in Eq.(2.2) è nullo. Questo è spesso vero nei problemi di Meccanica Celeste e fu alla base della descrizione del sistema del Mondo di Laplace e degli astronomi del XIX secolo.

Più in generale la difficoltà appena menzionata non si presenta se  $\vec{\omega}_0$  è tale che esistono due costanti  $C, \tau > 0$  per cui, posto  $|\vec{\nu}| \stackrel{def}{=} \max_{i=1, \dots, \ell} |\nu_i|$ ,

$$|\vec{\omega}_0 \cdot \vec{\nu}| > \frac{1}{C} |\vec{\nu}|^{-\tau} \quad \forall \vec{\nu} \in Z^\ell, \quad |\vec{\nu}| \neq 0 \quad (2.3)$$

detta “*proprietà Diofantea*”: invero i coefficienti di Fourier di una funzione analitica di  $\vec{\alpha}$  tendono a 0, per  $|\vec{\nu}| \rightarrow \infty$ , esponenzialmente e quindi le serie Eq.(2.2) certamente convergono.

*Osservazioni:* (1) La deduzione della Eq.(2.2) si basa solo sulla proprietà che  $\vec{A}(t) - \vec{A}(0) = O(\varepsilon)$  per  $0 \leq t \leq T$ .

(2) Va sottolineato che le relazioni ora ottenute sono relative al solo dato iniziale  $(\vec{A}_0, \vec{\alpha}_0)$  e conviene notare per riferimento futuro che il termine lineare in  $t$  può essere scritto, in termini della funzione

$$\overline{\Phi}(\vec{A}', \vec{\alpha}) \stackrel{def}{=} - \sum_{\vec{\nu} \neq \vec{0}} \frac{f_{\vec{\nu}}(\vec{A}')}{i\vec{\omega}_0 \cdot \vec{\nu}} e^{i\vec{\nu} \cdot \vec{\alpha}} \cdot \left(1 - \frac{((\vec{A}' - \vec{A}_0) \cdot \vec{\partial}_{\vec{A}'}) \vec{\omega}(\vec{A}') \cdot \vec{\nu}}{\vec{\omega}_0 \cdot \vec{\nu}}\right) \quad (2.4)$$

come  $\vec{\omega}' = (\vec{\omega}(\vec{A}'_0) + \varepsilon \vec{\partial}_{\vec{A}'} f_0(\vec{A}'_0)) t$  con  $\vec{A}'_0 = \vec{A}_0 - \varepsilon \vec{\partial}_{\vec{\alpha}} \overline{\Phi}(\vec{A}'_0, \vec{\alpha}_0)$  a meno di  $O(\varepsilon^2)$  e, inoltre, se

$$\vec{A}' = \vec{A} - \varepsilon \vec{\partial}_{\vec{\alpha}} \overline{\Phi}(\vec{A}', \vec{\alpha}), \quad \vec{\alpha}' = \vec{\alpha} + \varepsilon \vec{\partial}_{\vec{A}'} \overline{\Phi}(\vec{A}', \vec{\alpha}) \quad (2.5)$$

le Eq.(2.2) possono essere scritte anche

$$\vec{A}'(t) = \vec{A}'(0), \quad \vec{\alpha}'(t) = \vec{\alpha}'(0) + \vec{\omega}' t \quad (2.6)$$

(3) Nel precedente commento  $((\vec{A}' - \vec{A}'_0) \cdot \vec{\partial}_{\vec{A}'}) \vec{\omega}(\vec{A}')$  in  $\overline{\Phi}$  può essere sostituito da  $\vec{\omega}(\vec{A}') - \vec{\omega}(\vec{A}'_0)$ , a meno di  $O(\varepsilon^2)$ , *c.f.r.* §3.

### 3 Il teorema della media

Una formulazione alternativa, molto istruttiva e foriera di ulteriori sviluppi, del risultato perturbativo in §2 è il *teorema della media* che fornisce un formalismo Hamiltoniano per la descrizione dei moti con dati iniziali con azione

vicina entro  $O(1)$  a  $\vec{A}_0$ , anche se non più in termini di moti quasi periodici, tranne che per dati iniziali vicini entro  $O(\varepsilon)$  a  $\vec{A}_0$ , e comunque ancora con errore  $O(\varepsilon^2)$  o, per dati molto particolari, con errore  $O(\varepsilon^4)$  (per  $t \in [0, T]$ ).

Per limitare lo studio ai casi più semplici procederemo supponendo che  $\vec{\omega}_0$  soddisfi la proprietà Diofantea con costanti  $C, \tau$ . Si deve tener presente che la proprietà è vera per tutti gli  $\vec{\omega}_0 \in R^\ell$  tranne che per un insieme di  $\vec{\omega}_0$  a misura nulla in  $R^\ell$  (pur essendo una proprietà falsa per gli spettri  $\vec{\omega}_0$  a componenti razionali e quindi per un insieme pur denso in  $R^\ell$ ), *c.f.r.* Appendice A. Dunque definito

$$\Phi(\vec{A}', \vec{\alpha}) \stackrel{def}{=} - \sum_{\vec{\nu} \neq \vec{0}} \frac{f_{\vec{\nu}}(\vec{A}')}{i\vec{\omega}_0 \cdot \vec{\nu}} \left(1 - \frac{(\vec{\omega}(\vec{A}') - \vec{\omega}_0) \cdot \vec{\nu}}{\vec{\omega}_0 \cdot \vec{\nu}}\right) e^{i\vec{\nu} \cdot \vec{\alpha}} \quad (3.1)$$

e posto

$$\vec{A} = \vec{A}' + \varepsilon \vec{\partial}_{\vec{\alpha}} \Phi(\vec{A}', \vec{\alpha}), \quad \vec{\alpha}' = \vec{\alpha} + \varepsilon \vec{\partial}_{\vec{A}'} \Phi(\vec{A}', \vec{\alpha}) \quad (3.2)$$

si considerano i due cambiamenti di variabili (definiti nell'intorno di  $\{\vec{A}_0\} \times T^\ell$  via il teorema delle funzioni implicite) ottenuti, il primo, esprimendo  $\vec{\alpha}$  in funzione di  $\vec{\alpha}', \vec{A}'$  risolvendo la seconda delle Eq.(3.2) e sostituendo la soluzione  $\vec{\alpha} = \vec{\alpha}' + \varepsilon \vec{D}(\vec{A}', \vec{\alpha}')$  nella prima, e il secondo esprimendo  $\vec{A}'$  in funzione di  $\vec{A}, \vec{\alpha}$  risolvendo la prima delle Eq.(3.2) e sostituendo la soluzione  $\vec{A}' = \vec{A} + \varepsilon \vec{X}'(\vec{A}, \vec{\alpha})$  nella seconda.<sup>6</sup>

Scriveremo le due trasformazioni, che sono l'una inversa dell'altra, per costruzione, e inoltre sono *completamente canoniche* perché, derivando dalla Eq.(3.2), sono dotate di funzione generatrice  $\vec{A}' \cdot \vec{\alpha} + \varepsilon \Phi(\vec{A}', \vec{\alpha})$ , [1],

$$\begin{aligned} \vec{A}' &= \vec{A} + \varepsilon \vec{X}'(\vec{A}, \vec{\alpha}), & \vec{\alpha}' &= \vec{\alpha} + \varepsilon \vec{D}'(\vec{A}, \vec{\alpha}) \\ \vec{A} &= \vec{A}' + \varepsilon \vec{X}(\vec{A}', \vec{\alpha}'), & \vec{\alpha} &= \vec{\alpha}' + \varepsilon \vec{D}(\vec{A}', \vec{\alpha}') \end{aligned} \quad (3.3)$$

e la Eq.(2.2) si riscrive (trascurando sistematicamente quantità di  $O(\varepsilon^2)$  e facendo uso delle osservazioni seguenti la Eq.(2.3) e della Eq.(3.2))

---

<sup>6</sup>Una semplice scelta del dominio delle trasformazioni in Eq.(3.2) è che tali trasformazioni siano definite e analitiche su  $U' \times T^\ell$  e per  $|\varepsilon|$  piccolo se  $U' \subset U$  è un dominio chiuso con distanza  $> 0$  dalla frontiera di  $U$  e  $\varepsilon$  è abbastanza piccolo (in funzione della distanza di  $U'$  dalla frontiera di  $U$ ). Questo segue dall'analisi delle condizioni che si devono imporre per risolvere le equazioni implicite Eq.(3.2) facendo uso del teorema di Dini. L'analisi è elementare e non verrà discussa perché non essenziale in questi paragrafi introduttivi e non è usata nelle discussioni tecniche successive.

$$\begin{aligned}\vec{A}'(t) &= \vec{A}'(0) + O(\varepsilon^2), \\ \vec{\alpha}'(t) &= \vec{\alpha}'(0) + (\vec{\omega}(A'(0)) + \varepsilon \vec{\partial}_{\vec{A}'} f_{\vec{0}}(\vec{A}'(0)))t + O(\varepsilon^2)\end{aligned}\quad (3.4)$$

che descrive il moto con dati iniziali  $(\vec{A}_0, \vec{\alpha}_0)$  come un moto quasi periodico a meno di errore  $O(\varepsilon^2)$  nelle coordinate  $(\vec{A}', \vec{\alpha}')$ , con frequenze di rotazione *diverse da quelle del moto imperturbato*.

Se la matrice Hessiana  $\vec{\partial}_{\vec{A}\vec{A}}^2 h_0(\vec{A}) = \vec{\partial}_{\vec{A}} \vec{\omega}(\vec{A})$  ha determinante non nullo in  $\vec{A}_0$  allora l'equazione implicita per  $\vec{B}$ :  $\vec{\omega}(\vec{A}_0 + \vec{B}) + \varepsilon \vec{\partial}_{\vec{A}} f_{\vec{0}}(\vec{A}_0 + \vec{B}) = \vec{\omega}_0$ , può essere risolta se  $\varepsilon$  è abbastanza piccolo (via il teorema di Dini sulle funzioni implicite) ottenendo  $\vec{B}$  di  $O(\varepsilon)$  e il dato iniziale  $\vec{A}'_0 = \vec{A}_0 + \vec{B}$ ,  $\vec{\alpha}'(0)$  ha la notevole proprietà di evolvere a meno di  $O(\varepsilon^2)$  in modo che nelle nuove coordinate Eq.(3.3), entro  $O(\varepsilon^2)$ ,  $\vec{A}'$  è costante del moto e  $\vec{\alpha}'$  ruota con spettro  $\vec{\omega}_0$

$$\begin{aligned}\vec{A}'(t) &= \vec{A}'(0) + O(\varepsilon^2), \\ \vec{\alpha}'(t) &= \vec{\alpha}'(0) + \vec{\omega}_0 t + O(\varepsilon^2) = \vec{\alpha}'(0) + \vec{\partial}_{\vec{A}'} h_1(\vec{A}'(0))t + O(\varepsilon^2)\end{aligned}\quad (3.5)$$

ove  $h_1(\vec{A}') = h_0(\vec{A}') + \varepsilon f_{\vec{0}}(\vec{A}')$ .

Inoltre la trasformazione di coordinate Eq.(3.3) è una trasformazione completamente canonica perché dotata di funzione generatrice  $\Phi$ , c.f.r. Eq.(3.1), (3.2).

Quindi é naturale attendersi che l'Hamiltoniana  $H_\varepsilon$  in Eq.(1.2), se calcolata nelle coordinate  $(\vec{A}', \vec{\alpha}')$ , acquisisca una forma "semplice" o almeno tale che da essa si possa immediatamente dedurre che, se  $\vec{A}'$  è vicino entro  $O(\varepsilon)$  a  $\vec{A}_0$ , allora  $\vec{A}'$  è costante del moto e che gli angoli  $\vec{\alpha}'$  ruotano quasi periodicamente (tutto a meno di  $O(\varepsilon^2)$  e entro il tempo  $T$  fissato).

Lo si può verificare calcolando  $H_\varepsilon(\vec{A}' + \varepsilon \vec{X}'(\vec{A}', \vec{\alpha}'), \vec{\alpha}' + \varepsilon \vec{D}'(\vec{A}', \vec{\alpha}'))$ : per la validità simultanea delle Eq.(3.2) e (3.3) e aggiungendo e sottraendo opportune quantità, si trova

$$\begin{aligned}h(\vec{A}') + [h(\vec{A}' + \varepsilon \vec{\partial}_{\vec{\alpha}} \Phi(\vec{A}', \vec{\alpha})) - h(\vec{A}') - \varepsilon \vec{\omega}(\vec{A}') \cdot \vec{\partial}_{\vec{\alpha}} \Phi(\vec{A}', \vec{\alpha})] \\ + [\varepsilon \vec{\omega}(\vec{A}') \cdot \vec{\partial}_{\vec{\alpha}} \Phi(\vec{A}', \vec{\alpha}) + \varepsilon f(\vec{A}' + \varepsilon \vec{\partial}_{\vec{\alpha}} \Phi(\vec{A}', \vec{\alpha}), \vec{\alpha})]\end{aligned}\quad (3.6)$$

e poi si perviene, con un po' di passaggi algebrici riportati in dettaglio in Appendice A, ad un'interessante espressione dell'Hamiltoniana nelle nuove coordinate: tale espressione costituisce il *teorema della media* per una funzione di Hamilton, Eq.(1.2), analitica su  $U \times T^\ell$ :



**Teorema 1** (della media): Siano  $h(\vec{A})$  e  $f(\vec{A}, \vec{\alpha}, \varepsilon)$  funzioni analitiche nelle  $(\vec{A}, \vec{\alpha}, \varepsilon) \in U \times T^\ell \times \{|\varepsilon| < \varepsilon_0\}$  e sia  $\vec{\omega}_0 = \vec{\partial}_{\vec{A}} h(\vec{A}_0)$  uno spettro di frequenze Diofanteo, c.f.r. Eq.(2.3), e sia  $\neq 0$  il determinante Hessiano  $\det \vec{\partial}_{\vec{A}\vec{A}}^2 h(\vec{A}_0)$  (“condizione di torsione”). Allora prefissato un tempo  $T > 0$ :

(1) il sistema ammette coordinate canoniche  $(\vec{A}', \vec{\alpha}')$  in cui l’Hamiltoniana assume, se  $h_1(\vec{A}') \stackrel{def}{=} h(\vec{A}') + \varepsilon f_{\vec{0}}(\vec{A}')$ , la forma  $H_\varepsilon^1$ :

$$H_\varepsilon^1(\vec{A}', \vec{\alpha}') \stackrel{def}{=} h_1(\vec{A}') + \varepsilon^2 f_2(\vec{A}', \vec{\alpha}', \varepsilon) + \varepsilon(\vec{\omega}(\vec{A}') - \vec{\omega}(\vec{A}'_0)) \cdot M_2(\vec{A}', \vec{\alpha}', \varepsilon)(\vec{\omega}(\vec{A}') - \vec{\omega}(\vec{A}'_0)) \quad (3.7)$$

con  $\vec{A}'_0$  opportuno, vicino a  $\vec{A}_0$  entro  $O(\varepsilon)$  e tale che  $\vec{\partial}_{\vec{A}'} h_1(\vec{A}'_0) = \vec{\omega}_0$ , e con  $M_2, f_2$ , rispettivamente matrice  $\ell \times \ell$  e funzione scalare, analitiche in  $(\vec{A}, \vec{\alpha}, \varepsilon) \in U' \times T^\ell \times \{|\varepsilon| < \varepsilon'_0\}$  con  $U' \subset U$  chiuso, a distanza  $\lambda$  arbitrariamente prefissata  $> 0$  dalla frontiera di  $U$ , e  $\varepsilon'_0$  opportunamente piccolo (dipendente da  $\lambda$ , c.f.r. nota<sup>5</sup>).

(2) Quindi, su tempi  $t \in [0, T]$  e per tutti i dati iniziali con  $|\vec{A}' - \vec{A}'_0| = O(\varepsilon)$ , i moti sono quasi periodici a meno di  $O(\varepsilon^2)$  e i moti con dato iniziale  $\vec{A}'_0$  hanno spettro di rotazione  $\vec{\omega}_0$ .

(3) Le coordinate  $(\vec{A}', \vec{\alpha}')$  in cui il moto è ben descritto a meno di  $O(\varepsilon^2)$  sono esplicitamente determinate dalle Eq.(3.2).

(4) Se  $f$  è un polinomio trigonometrico il risultato vale nell’ipotesi pi’u debole che  $\vec{\omega}_0 \cdot \vec{\nu} \neq 0$  almeno per tutte le armoniche  $\vec{\nu}$  per cui  $f_{\vec{\nu}}(\vec{A}) \neq 0$  in un intorno di  $\vec{A}_0$ .

*Osservazioni:* (a) Esaminando la deduzione (in Appendice A) della  $H_\varepsilon^1(\vec{A}', \vec{\alpha}')$  si nota che se in luogo della  $\Phi$  in Eq.(3.1) si fosse usata la funzione  $\tilde{\Phi}(\vec{A}', \vec{\alpha}') \stackrel{def}{=} -\sum_{\vec{\nu} \neq \vec{0}} \frac{1}{i\vec{\omega} \cdot \vec{\nu}} f_{\vec{\nu}} e^{i\vec{\nu} \cdot \vec{\alpha}}$  si sarebbe ottenuto un teorema molto simile con  $H_\varepsilon^1$  della forma  $h_1(\vec{A}') + \varepsilon(\vec{A}' - \vec{A}'_0) \cdot \vec{\varphi}_2(\vec{A}', \vec{\alpha}') + O(\varepsilon^2)$ : in tal caso non si potrebbe affermare che i moti con dato iniziale  $\vec{A}'$  vicino entro  $O(\varepsilon)$  a  $\vec{A}'_0$  sono quasi periodici (per  $0 \leq t \leq T$ ) con precisione  $O(\varepsilon^2)$ : questo perché  $\vec{\alpha}'$  non è costante già all’ordine  $\varepsilon$ . Quindi il risultato darebbe notevolmente meno informazioni di quello del teorema della media, che invece dà quasi periodicità approssimata entro  $O(\varepsilon^2)$  a dati iniziali  $\vec{A}'$  vicini entro  $O(\varepsilon)$  a  $\vec{A}'_0$ .

(b) Inoltre la proporzionalità a  $(\vec{A}' - \vec{A}'_0)^2$  del termine di ordine  $\varepsilon$  in Eq.(3.7) implica che se  $\vec{A}'(0) = \vec{A}'_0$  allora  $\vec{A}'(t) - \vec{A}'_0$  rimane di ordine  $O(\varepsilon^2)$  per  $0 \leq t \leq T$  e, ripetendo la discussione del §2 (con  $(\vec{A}', \vec{\alpha}')$  che giocano il ruolo di  $(\vec{A}, \vec{\alpha})$

e con nuove coordinate  $(\vec{A}'', \vec{\alpha}'')$  che giocano il ruolo di  $(\vec{A}', \vec{\alpha}')$  si perviene all'esistenza di dati iniziali  $(\vec{A}_0'', \vec{\alpha}'')$  che evolvono quasi periodicamente con spettro  $\vec{\omega}_0$  a meno di  $O(\varepsilon^4)$ . Ma questo è notevolmente migliorato dal teorema del seguente §4, e quindi si tralasciano ulteriori dettagli.

#### 4 Teorema della media di Kolmogorov

Un diverso teorema della media è basato su un altro punto di vista: invece di cercare coordinate  $(\vec{A}', \vec{\alpha}')$  in cui l'Hamiltoniana divenga una perturbazione di ordine  $\varepsilon^2$  di un sistema integrabile si cerca di trasformarla in una che differisca per quantità di ordine cubico nelle azioni  $\vec{A}' - \vec{A}'_0$  (eventualmente anche solo di ordine  $\varepsilon$ ) e a meno di  $O(\varepsilon^2)$  da un'Hamiltoniana “quadratica” (cioè con soli termini lineari e quadratici) nelle azioni imponendo che quest'ultima, pur non essendo integrabile, ammetta, in assenza di perturbazioni, rigorosamente un toro invariante percorso quasi periodicamente con vettore di rotazione  $\vec{\omega}_0$  per  $\vec{A}' = \vec{A}'_0$ .

Il vantaggio *notevolissimo* è che il procedimento può essere “facilmente” iterato arrivando dopo  $k$  iterazioni a coordinate  $(\vec{A}', \vec{\alpha}')$  in cui l'Hamiltoniana differisce da un'Hamiltoniana quadratica nelle azioni solo a meno di termini di ordine cubico nelle azioni  $\vec{A}' - \vec{A}'_0$  a meno di una funzione  $\sim O(\varepsilon^{2k})$ , pur mantenendo il dominio di definizione  $U' \times T^\ell$  con  $U'$  contenente una sfera con raggio di  $O(1)$  e centro  $\vec{A}'_0$ .

L'idea è basata sul *ruolo speciale* delle trasformazioni canoniche con funzione generatrice  $\vec{B} \cdot \vec{\alpha} + X(\vec{\alpha}) + \vec{A}' \cdot \vec{Y}(\vec{\alpha})$  “lineare nelle azioni”, con  $\vec{B}$  vettore costante,  $X, \vec{Y}$  funzioni rispettivamente scalare e vettoriale. Queste trasformazioni hanno la forma molto particolare:

$$\vec{A} = \vec{B} + \vec{\partial}_{\vec{\alpha}} X(\vec{\alpha}) + \vec{\partial}_{\vec{\alpha}}(\vec{A}' \cdot \vec{Y}(\vec{\alpha})), \quad \vec{\alpha}' = \vec{\alpha} + \vec{Y}(\vec{\alpha}) \quad (4.1)$$

Allora se  $H(\vec{A}, \vec{\alpha}) = \vec{\omega}_0 \cdot \vec{A} + \frac{1}{2} \vec{A} \cdot M(\vec{\alpha}) \vec{A} + (f(\vec{\alpha}) + \vec{A} \cdot \vec{g}(\vec{\alpha}))$  le trasformazioni invertibili con funzione generatrice lineare in  $\vec{A}'$  trasformano  $H(\vec{A}, \vec{\alpha})$  in  $H'(\vec{A}', \vec{\alpha}')$  che ha la stessa forma, ovviamente con nuove funzioni  $M', f', \vec{g}'$ : ossia le Hamiltoniane di questa forma sono una famiglia *invariante rispetto a trasformazioni canoniche con generatrici lineari nelle azioni*.

Supponendo, senza perdita di generalità,  $\vec{A}'_0 = \vec{0}$  si consideri l'Hamiltoniana  $h(\vec{A}) + \varepsilon f(\vec{A}, \vec{\alpha}, \varepsilon)$  con  $f(\vec{A}, \vec{\alpha}, \varepsilon) = f(\vec{A}, \vec{\alpha})$  indipendente, per semplicità notazionale, da  $\varepsilon$ . Nell'intorno di  $\vec{A}'_0$  la si può riscrivere, sviluppando sia  $h$

che  $f$  in serie di Taylor nelle  $\vec{A}$ , intorno ad  $\vec{A} = \vec{A}_0 = \vec{0}$  fino al secondo ordine incluso, come

$$H_\varepsilon(\vec{A}, \vec{\alpha}) = h(\vec{0}) + \vec{\omega}_0 \cdot \vec{A} + \frac{1}{2} \vec{A} \cdot (M + m(\vec{\alpha})) \vec{A} + \varepsilon(\varphi(\vec{\alpha}) + \vec{A} \cdot \vec{\gamma}(\vec{\alpha})) + O(\vec{A}^3) \quad (4.2)$$

ove:  $m = \varepsilon\mu$ ,  $M = \partial_{\vec{A}\vec{A}}^2 h(\vec{0})$  è la matrice Hessiana di  $h$ , e  $\mu, \vec{\gamma}, \varphi$  sono funzioni (a valori matrice  $\ell \times \ell$ , vettore in  $R^\ell$  e scalare rispettivamente) legate alle prime due derivate in  $\vec{A}$  di  $f$ .

Si reimposti allora la deduzione di un nuovo teorema della media proponendo la funzione generatrice  $\vec{A}' \cdot \vec{\alpha} + \varepsilon\Psi(\vec{A}', \vec{\alpha})$  con  $\Psi(\vec{A}', \vec{\alpha}) = \vec{b} \cdot \vec{\alpha} + X(\vec{\alpha}) + \vec{A}' \cdot \vec{Y}(\vec{\alpha})$  e con  $X$  e  $\vec{Y}$  scelti in modo che

$$\begin{aligned} (\vec{\omega}_0 \cdot \vec{\partial}_{\vec{\alpha}}) X(\vec{\alpha}) + \varphi(\vec{\alpha}) &= 0 \\ (\vec{\omega}_0 \cdot \vec{\partial}_{\vec{\alpha}}) \vec{Y}(\vec{\alpha}) + (M + m(\vec{\alpha}))(\vec{b} + \vec{\partial}_{\vec{\alpha}} X(\vec{\alpha})) + \vec{\gamma}(\vec{\alpha}) &= \vec{0} \end{aligned} \quad (4.3)$$

Le funzioni  $X, \vec{Y}$  si calcolano per trasformata di Fourier se  $\vec{b}$  è tale che la media su  $\vec{\alpha}$  di  $(M + m(\vec{\alpha}))(\vec{b} + \vec{\partial}_{\vec{\alpha}} X(\vec{\alpha})) + \vec{\gamma}$  è nulla e si trova, quindi, se  $M + m_{\vec{0}}$  è invertibile:

$$\begin{aligned} \vec{b} &\stackrel{def}{=} - (M + m_{\vec{0}})^{-1} [\vec{\gamma}_{\vec{0}} + \sum_{\vec{v} \neq \vec{0}} m_{-\vec{v}} \frac{-i\vec{v} \cdot \varphi_{\vec{v}}}{i\vec{\omega}_0 \cdot \vec{v}}] \\ \Psi(\vec{A}', \vec{\alpha}) &\stackrel{def}{=} \vec{b} \cdot \vec{\alpha} - \sum_{\vec{v} \neq \vec{0}} \left( \frac{\varphi_{\vec{v}} + \vec{A}' \cdot \vec{\gamma}_{\vec{v}}}{i\vec{\omega}_0 \cdot \vec{v}} - \frac{\vec{A}' \cdot M \vec{v} \varphi_{\vec{v}}}{i(\vec{\omega}_0 \cdot \vec{v})^2} \right) e^{i\vec{v} \cdot \vec{\alpha}} \\ &\quad - \sum_{\vec{v} \neq \vec{0}} \frac{1}{i\vec{\omega}_0 \cdot \vec{v}} \vec{A}' \cdot \left( m_{\vec{v}} \vec{b} - \sum_{\substack{\vec{v}' + \vec{v}'' = \vec{v} \\ \vec{v}', \vec{v}'' \neq \vec{0}}} m_{\vec{v}'} i\vec{v}'' \frac{\varphi_{\vec{v}''}}{i\vec{\omega}_0 \cdot \vec{v}''} \right) e^{i\vec{v} \cdot \vec{\alpha}} \end{aligned} \quad (4.4)$$

ove  $m_{\vec{0}}$  è la media di  $m(\vec{\alpha})$  su  $\vec{\alpha}$ : si noti che se  $m_{\vec{v}} = 0$  per tutti i  $\vec{v} \neq \vec{0}$  la  $\Psi$  è somma di  $\vec{b} \cdot \vec{\alpha}$  e della “parte lineare” in  $\vec{A}'$  della Eq.(3.1).

Se la matrice  $m$  ha componenti di Fourier con  $\vec{v} \neq \vec{0}$  non nulle (ossia se  $m$  non è costante) e di  $O(1)$ , la  $\vec{A}' \cdot \vec{\alpha} + \varepsilon\Psi$  genera una trasformazione canonica che differisce  $O(\varepsilon)$ , cioè sostanzialmente, dalla corrispondente considerata nel teorema della media: e ripercorrendo la dimostrazione del teorema 1 applicandola alla Hamiltoniana Eq.(4.2), ossia sempre alla  $h + \varepsilon f$  della Eq.(1.2), si trova come risultato

$$\begin{aligned}
H'(\vec{A}', \vec{\alpha}') = & h(\vec{0}) + \vec{\omega}_0 \cdot \vec{A}' + \frac{1}{2} \vec{A}' \cdot (M + m_2(\vec{\alpha}')) \vec{A}' \\
& + \varepsilon^2 (\varphi_2(\vec{\alpha}') + \vec{A}' \cdot \vec{\gamma}_2(\vec{\alpha}')) + O(\vec{A}'^3)
\end{aligned} \tag{4.5}$$

con  $m_2, \vec{\gamma}_2, \varphi_2$  opportune funzioni analitiche di  $\vec{\alpha}'$  e  $\varepsilon$  e  $m_2 = \varepsilon\mu + O(\varepsilon^2)$ .

In questa verifica, di natura puramente algebrica, è importante notare che è solo necessario che la matrice  $M + m(\vec{\alpha})$  abbia media  $M + m_{\vec{0}}$  invertibile: (il che per la Eq.(4.2) è certo vero per  $\varepsilon$  piccolo perché  $m(\vec{\alpha}) = \varepsilon\mu(\vec{\alpha})$ ). Quindi alla stessa conclusione si perviene se la matrice  $m(\vec{\alpha})$  in Eq.(4.2) è  $m_1(\vec{\alpha}) + \varepsilon\mu(\vec{\alpha})$  con  $m_1 = O(1)$  e  $M + m_{1\vec{0}}$  invertibile.

**Teorema 2:** (*Kolmogorov*) sia  $\vec{\omega}_0$  uno spettro di frequenze Diofanteo, c.f.r. Eq.(2.3), e si consideri la Hamiltoniana in Eq.(4.2) con  $m(\vec{\alpha}) = m_1(\vec{\alpha}) + \varepsilon\mu(\vec{\alpha})$  matrice  $\ell \times \ell$  analitica in  $\vec{\alpha}, \varepsilon$  (anche se di  $O(1)$ ) soggetta alla sola condizione che  $M + m_{1\vec{0}}$  sia invertibile

(1) il sistema ammette coordinate canoniche  $(\vec{A}', \vec{\alpha}')$  in cui la Hamiltoniana assume la forma Eq.(4.5) e con  $\vec{\gamma}_2, \varphi_2, m_2$  analitiche in  $(\vec{\alpha}', \varepsilon) \in U'' \times T^\ell \times \{|\varepsilon| < \varepsilon'_0\}$  con  $U''$  intorno di raggio  $\rho$  abbastanza piccolo di  $\vec{0}$  e  $m_2(\vec{\alpha}') = m_1(\vec{\alpha}') + O(\varepsilon)$ .<sup>7</sup>

(2) le coordinate  $(\vec{A}', \vec{\alpha}')$  in cui il moto è ben descritto a meno di  $O(\vec{A}'^3)$  sono esplicitamente determinate dalla funzione generatrice Eq.(4.4).

Si noti che a differenza del teorema della media questo risultato non può, in generale, essere visto come una teoria perturbativa di un sistema integrabile ma piuttosto come una teoria perturbativa del sistema di Hamiltoniana  $H(\vec{A}, \vec{\alpha}) = \vec{\omega}_0 \cdot \vec{A} + \frac{1}{2} \vec{A} \cdot (M + m_1(\vec{\alpha})) \vec{A} + O(\vec{A}^3)$  vicino a  $\vec{A} = \vec{0}$ . Tale  $H$  non è integrabile: tuttavia possiede un moto esattamente quasi periodico sul toro  $\{\vec{0}\} \times T^\ell$ . Solo se applicato al caso  $m = 0$  dà lo stesso risultato del teorema della media.

## 5 Osservazioni sui teoremi della media

(1) Il nome “teorema della media”, teorema 1, è dovuto alla modifica dell’Hamiltoniana integrabile da  $h(\vec{A})$  ad  $h_1(\vec{A})$  somma di  $h(\vec{A})$  e della media sugli angoli,  $\varepsilon f_{\vec{0}}(\vec{A}) \stackrel{def}{=} \varepsilon \int_{T^\ell} f(\vec{A}, \vec{\alpha}) \frac{d\vec{\alpha}}{(2\pi)^\ell}$ , della perturbazione  $\varepsilon f$ .

(2) Dunque il dato iniziale  $(\vec{A}_0, \vec{\alpha}_0)$  evolve, entro il tempo prefissato e a

---

<sup>7</sup>Si noti:  $O(\varepsilon)$  e non  $O(\varepsilon^2)$ .

meno di  $O(\varepsilon^2)$ , in modo quasi periodico ma con uno spettro di frequenze  $\vec{\omega}'_0 \stackrel{def}{=} (\vec{\omega}_0 + \varepsilon \vec{\partial}_{\vec{A}} f_0(\vec{A}_0) + \varepsilon \sum_{\vec{\nu} \neq \vec{0}} \frac{f_{\vec{\nu}}(\vec{A}_0) i(\vec{\nu} \cdot \vec{\partial}_{\vec{A}}) \vec{\omega}(\vec{A}_0)}{i\vec{\omega}_0 \cdot \vec{\nu}} e^{i\vec{\nu} \cdot \vec{\alpha}_0})$  diverso da quello imperturbato  $\vec{\omega}_0$ , mentre  $O(\varepsilon)$  vicino a esso ci sono altri dati iniziali che dopo la perturbazione si muovono con lo stesso spettro e, novità del teorema 1 rispetto al calcolo del §2, entro  $O(\varepsilon^4)$ , c.f.r. osservazione (b) al teorema 1.

(3) Nelle coordinate  $(\vec{A}, \vec{\alpha})$  il moto con dato iniziale  $(\vec{A}_0, \vec{\alpha}_0)$  evolve con spettro  $\vec{\omega}_0$  a meno di  $O(\varepsilon)$ , mentre nelle coordinate  $(\vec{A}', \vec{\alpha}')$  evolve con spettro diverso  $\vec{\omega}'_0$  a meno di  $O(\varepsilon^2)$ : questo è dovuto alla limitatezza del tempo  $T$  di osservazione e alla differenza di  $O(\varepsilon)$  fra  $\vec{A}'$  e  $\vec{A}$ ; ma si noti che un moto solo approssimativamente quasi periodico non ha spettro definito senza ambiguità.

(4) Oltre a questi ultimi dati iniziali anche i dati iniziali con  $\vec{A} - \vec{A}_0 = O(\varepsilon)$  evolvono (entro  $T$ ) in modo quasi periodico con errore  $O(\varepsilon^2)$ . Tuttavia il teorema 1 vale in un intorno  $U'$  di ordine  $O(1)$ , se  $\varepsilon$  è abbastanza piccolo. Quindi la maggior parte dei dati iniziali evolve ancora solo in modo quasi periodico con errore grande  $O(\varepsilon)$  e non  $O(\varepsilon^2)$ .

(5) Sarebbe errato pensare che l'errore sul moto con dato iniziale con  $\vec{A}' - \vec{A}'_0 = O(\varepsilon)$  si possa stimare, via il teorema della media, teorema 1, entro un tempo arbitrario  $t$  come  $O(\varepsilon^2 t)$ : infatti ogni tentativo di stimare l'errore nella Eq.(2.1) porta inevitabilmente, in generale, a stime di errore non migliori di  $O(\varepsilon(e^{O(\varepsilon)t} - 1))$  e quindi per tempi  $t \simeq O(\varepsilon^{-1})$  la stima degli errori, come sopra studiata, è illusoria. Per applicare il teorema della media si deve prima fissare "il tempo di osservazione  $T$ " e poi vedere se  $\varepsilon$  è abbastanza piccolo perché valgano le conclusioni del teorema.

(6) Il teorema della media, teorema 1, mostra la vera difficoltà del problema che si presenterebbe se si volesse ottenere una descrizione del moto in approssimazioni  $\ll O(\varepsilon)$  per tutti i dati con  $\vec{A}$  nell'intorno di  $\vec{A}_0$ . Anche quando si può applicare il teorema della media, o perché la perturbazione è un polinomio trigonometrico con armoniche  $\vec{\nu}$  non risonanti con  $\vec{\omega}_0$  ovvero perché  $\vec{\omega}_0$  ha la proprietà Diofantea Eq.(2.3), si perviene ad una descrizione dei moti, in nuove coordinate, con frequenze di rotazione un po' diverse e con errore ancora di ordine  $O(\varepsilon)$  tranne che per dati iniziali  $\vec{A}$  vicini entro  $O(\varepsilon)$  a  $\vec{A}_0$  e quindi a  $\vec{A}'_0$ .

Come si vede dalla (3) dell'Appendice A ove si svolgono i passaggi algebrici che conducono al teorema 1, per eliminare il termine di  $O(\varepsilon)$  e quadratico in  $\vec{A}' - \vec{A}'_0$ , che appare nella Eq.(3.7) conclusiva, la  $\Phi$  dovrebbe risolvere

$\vec{\omega}(\vec{A}') \cdot \vec{\partial}_{\vec{\alpha}} \Phi(\vec{A}', \vec{\alpha}) = -(f(\vec{A}', \vec{\alpha}) - f_0(\vec{A}'))$  e cioè

$$\Phi(\vec{A}', \vec{\alpha}) = - \sum_{\vec{v} \neq \vec{0}} \frac{f_{\vec{v}}(\vec{A}')}{i\vec{\omega}(\vec{A}') \cdot \vec{v}} e^{i\vec{v} \cdot \vec{\alpha}} \quad (5.1)$$

che però richiede che  $\vec{\omega}(\vec{A}') \cdot \vec{v}$  non sia mai nullo nelle vicinanze di  $\vec{A}'$  il che non è in generale possibile se la matrice Hessiana di  $h$  ha determinante diverso da 0 (perché, come già notato, questo implica che al variare di  $\vec{A}'$  in un intorno di  $\vec{A}_0$  lo spettro  $\vec{\omega}(\vec{A}')$  va in un intorno di  $\vec{A}_0$ ).

Si noti, in proposito, che la scelta della  $\Phi$  in Eq.(3.1) si ottiene formalmente dalla Eq.(5.1) sviluppando  $(\vec{\omega}(\vec{A}') \cdot \vec{v})^{-1} \equiv (\vec{\omega}_0 \cdot \vec{v} + (\vec{\omega}(\vec{A}') - \vec{\omega}_0) \cdot \vec{v})^{-1}$  in serie di  $\frac{(\vec{\omega}(\vec{A}') - \vec{\omega}_0) \cdot \vec{v}}{\vec{\omega}_0 \cdot \vec{v}}$  al primo ordine.

(7) Fa eccezione il caso in cui  $f$  è un polinomio trigonometrico e  $\vec{\omega}_0 \cdot \vec{v} \neq 0$  per tutte le armoniche  $\vec{v}$  presenti in  $f$ : allora la Eq.(5.1) è ben definita vicino a  $\vec{A}_0$  e nella Eq.(3.7) viene a mancare il termine di  $O(\varepsilon)$ . Tuttavia in tal caso la funzione  $f_2$ , in generale, non sarà più un polinomio trigonometrico e la difficoltà si ripresenterà tentando di ottenere, riapplicando il teorema della media alla Eq.(3.7), un'approssimazione di  $O(\varepsilon^4)$ . Dunque nasce la possibilità che tentando di migliorare ulteriormente l'approssimazione il nuovo spettro di frequenze  $\vec{\omega}'(\vec{A}') = \vec{\omega}(\vec{A}') + \varepsilon \vec{\partial}_{\vec{A}'} f_0(\vec{A}')$  sia risonante e non si possa procedere a migliorare l'approssimazione: e il problema fu posto in luce e in forma precisa da Poincaré che mostrò come effettivamente si presentasse nei più semplici problemi di Meccanica Celeste, [6].

(8) L'osservazione (2) mostra che se  $\vec{\omega}_0 = \vec{\omega}(\vec{A}_0)$  è Diofanteo, fra i moti imperturbati se ne può trovare uno che, entro  $O(\varepsilon^4)$  e per  $t \in [0, T]$ , è quasi periodico con spettro  $\vec{\omega}_0$ , almeno se la matrice Hessiana di  $h$  è non degenera. Allora diventa concepibile pensare di poter dimostrare l'esistenza di un moto quasi periodico con spettro  $\vec{\omega}_0$  adattando la costruzione che conduce al teorema della media, teorema 1, in modo che si possa iterare indefinitamente. Però una iterazione ingenua non è possibile perché l'Hamiltoniana che si deduce dal teorema della media *non ha la forma iniziale* in quanto alcuni termini sono di  $O(\varepsilon^2)$  se  $\vec{A}' - \vec{A}'_0 \sim O(\varepsilon)$  ma formalmente sono solo di ordine  $\varepsilon$  (c.f.r. Eq.(3.7)) e quindi  $H_\varepsilon^1$  ha struttura diversa da quella iniziale di  $H_\varepsilon$  e per mantenerla può essere necessario restringere il dominio di definizione di  $H_\varepsilon^1$ : *a priori* talmente da rendere difficile mostrare che  $\varepsilon^2 f_2$  è effettivamente di  $O(\varepsilon^2)$ . Con ulteriori idee questa strada è tuttavia perseguibile come mostrato successivamente da Arnold, [4, 7].

La corretta impostazione del problema dell'esistenza di un moto quasi periodico di spettro  $\vec{\omega}_0$  sulla base del commento (8) fu proposta da Kolmogorov che indicò la possibilità, facendo uso non del teorema 1 della media ma del teorema 2, di mantenere il dominio di analiticità della Hamiltoniana della forma  $U_0 \times T^\ell \times \{|\varepsilon| < \varepsilon_K\}$  con  $U_0$  una sferetta con centro  $\vec{A}_0^\infty$  opportuno e  $\varepsilon_K > 0$  abbastanza piccolo. E svolgere in questo infinite iterazioni pervenendo a definire coordinate analitiche e canoniche  $\vec{A}^\infty, \vec{\alpha}^\infty$  in cui l'Hamiltoniana originale è trasformata *esattamente* in

$$H_\varepsilon^\infty = \vec{\omega}_0 \cdot \vec{A}^\infty + \frac{1}{2}(\vec{A}^\infty - \vec{A}_0^\infty) \cdot M_\infty(\vec{\alpha}^\infty, \varepsilon)(\vec{A}^\infty - \vec{A}_0^\infty), \quad (5.2)$$

$$+ O((\vec{A}^\infty - \vec{A}_0^\infty)^3)$$

Tale Hamiltoniana genera evidentemente un moto *esattamente* quasi periodico di spettro  $\vec{\omega}_0$  per  $\vec{A}^\infty = \vec{A}_0^\infty, \vec{\alpha}^\infty$  arbitrario e per  $t \leq T$  (e quindi *per tutti i tempi* per l'arbitrarietà di  $\vec{\alpha}^\infty$ ), ma non dà informazioni sui moti generati da dati vicini ove la perturbazione resta di  $O(\varepsilon)$  a causa della dipendenza da  $\varepsilon$  della matrice  $M(\vec{\alpha}^\infty, \varepsilon)$ . È questo il contenuto del *teorema KAM* nella versione di Kolmogorov, [3].

Quest'affermazione è facilmente deducibile dal breve testo originale: specialmente (*c.f.r.* Eq.(4.1)) nel caso

$$H_\varepsilon(\vec{A}, \vec{\alpha}) = \vec{\omega}_0 \cdot \vec{A} + \frac{1}{2J} \vec{A} \cdot \vec{A} + \varepsilon(f(\vec{\alpha}) + \vec{A} \cdot \vec{g}(\vec{\alpha}) + \vec{A} \cdot m(\vec{\alpha})\vec{A}) \quad (5.3)$$

con  $J > 0$  e  $f, \vec{g}, m$  uno scalare, un vettore e una matrice funzioni analitiche o in casi marginalmente più generali. E allora la dimostrazione del teorema di Kolmogorov si può iterare.

Il procedimento di Kolmogorov è *a posteriori* molto naturale e semplice, anche nel caso generale (ossia Eq.(1.2),  $M = \vec{\partial}_{\vec{A}\vec{A}}^2 h(\vec{A}_0)$  invertibile e  $\vec{\omega}_0$  Diofanteo): ma, oltre al nuovo teorema del §4, certo richiede comunque un controllo degli infiniti cambiamenti di variabile e un'analisi dettagliata dell'*accumulazione degli errori*: Kolmogorov si limitò al solo primo passo dell'iterazione, ossia al teorema 2, (*correttamente*) concludendo: “*The construction of further approximations is not associated with new difficulties*”.

Ma a causa proprio dell'originalità dell'idea furono necessari molti anni ( $\sim 5$  lustri) prima che il risultato di Kolmogorov venisse universalmente accettato come rigoroso e completo: a questo contribuì in particolare [8], mentre fino ad allora molti avevano ritenuto che il teorema di Kolmogorov fosse

quello esposto in dettaglio da Arnold in [4], che però seguì un procedimento piuttosto diverso.

É anche possibile un altro punto di vista, altrettanto naturale, che si propone di determinare il moto quasi periodico di spettro  $\vec{\omega}_0$  semplicemente generando i coefficienti di Taylor in  $\varepsilon$  delle equazioni parametriche del toro su cui ha luogo e verificare la convergenza delle serie relative. La possibilità della determinazione univoca di questi coefficienti era nota dal XIX secolo (serie di Lindstedt e Newcomb) ma questa idea fu inizialmente poco seguita per la forte influenza del lavoro di Kolmogorov (come diffuso nella versione di Arnold). I metodi di Arnold e Kolmogorov sono ormai esposti ampiamente e, anche per questo, sarà utile illustrare nei prossimi paragrafi questo procedimento, interessante per la sua natura squisitamente combinatoria, alternativo sia a [3] che a [4]: ci limiteremo per semplicità a un caso in cui l'Hamiltoniana ha forma Eq.(5.3) invariante per le trasformazioni lineari nelle azioni e precisamente all'Hamiltoniana in Eq.(6.2), seguente.

## 6 Teorema KAM

Il risultato generale è il teorema cardine della teoria della stabilità Hamiltoniana per sistemi quasi integrabili:

**Teorema 3:** (KAM) *Nelle ipotesi di analiticità su  $h, f$ , siano  $\vec{\omega}_0 \stackrel{def}{=} \vec{\omega}(\vec{A}_0)$  le velocità angolari dei moti imperturbati generati sul toro  $\{\vec{A}_0\} \times T^\ell$  dall'Hamiltoniana  $h$ . Si supponga che  $\vec{\omega}_0$  verifichi una disuguaglianza Diofantea, c.f.r. Eq.(2.3) e, inoltre, che la matrice Hessiana  $\partial_{\vec{A}\vec{A}}^2 h(\vec{A})$  di  $h$  sia non degenere. Allora esistono due funzioni  $\vec{\Xi}_\varepsilon(\vec{\psi}), \vec{\Delta}_\varepsilon(\vec{\psi})$ , analitiche in  $\vec{\psi} \in T^\ell$  e in  $\varepsilon$  per  $|\varepsilon| < \varepsilon_0$  con  $\varepsilon_0$  abbastanza piccolo, tali che*

$$\vec{A}(t) = \vec{A}_0 + \varepsilon \vec{\Xi}_\varepsilon(\vec{\psi} + \vec{\omega}_0 t), \quad \vec{\alpha}(t) = \vec{\psi} + \vec{\omega}_0 t + \varepsilon \vec{\Delta}_\varepsilon(\vec{\psi} + \vec{\omega}_0 t) \quad (6.1)$$

risolvono le equazioni del moto per  $H_\varepsilon \stackrel{def}{=} h(\vec{A}) + \varepsilon f(\vec{A}, \vec{\alpha}, \varepsilon)$ .

Quindi i moti quasi periodici imperturbati con rotazione  $\vec{\omega}_0$  Diofantea si “ritrovano” con le stesse frequenze di rotazione fra i moti generati dall'Hamiltoniana perturbata e si svolgono su tori invarianti leggermente deformati, di equazioni parametriche  $\vec{A} = \vec{A}_0 + \varepsilon \vec{\Xi}_\varepsilon(\vec{\psi})$ ,  $\vec{\alpha} = \vec{\psi} + \varepsilon \vec{\Delta}_\varepsilon(\vec{\psi})$ ,  $\vec{\psi} \in T^\ell$ .

Già per “ $\ell$  rotatori” interagenti con potenziale polinomiale  $P(\vec{\alpha})$  pari” e



indipendente da  $\varepsilon$

$$h(\vec{A}) = \frac{1}{2J} \vec{A} \cdot \vec{A}, \quad f(\vec{A}, \vec{\alpha}) \equiv P(\vec{\alpha}) = \sum_{|\vec{\nu}| < N} f_{\vec{\nu}} e^{i\vec{\nu} \cdot \vec{\alpha}}, \quad f_{\vec{\nu}} = f_{-\vec{\nu}} \in R \quad (6.2)$$

con  $J > 0$  e  $N < \infty$ , si ha un caso che contiene tutte le difficoltà essenziali del caso più generale in cui la matrice Hessiana di  $h$  è non degenere e  $f$  è semplicemente una funzione analitica.

Un risultato importante sulla *convergenza* delle serie perturbative fu ottenuto da Siegel, [2], risolvendo un problema sulla linearizzabilità di trasformazioni olomorfe di  $C^\ell$ , più semplice ma strettamente collegato al problema dei moti quasi periodici Hamiltoniani, e aprì la strada al superamento della simile, secolare, “difficoltà dei piccoli denominatori” in meccanica Hamiltoniana. Furono, però, necessari nove lustri perché l’idea fosse ripresa e applicata per ottenere il teorema KAM e un altro perché il lavoro relativo fosse pubblicato, [9].

Qui esporremo la estensione del metodo di Siegel alla dimostrazione del teorema KAM nel caso Eq.(6.2). Le linee della dimostrazione saranno:

(1) Costruzione delle serie di potenze in  $\varepsilon$  per le funzioni  $\vec{\Xi}_\varepsilon, \vec{\Delta}_\varepsilon$  imponendo che queste funzioni risolvano le equazioni del moto: tali serie sono note come *serie di Lindstedt*; la costruzione prescinde dalla convergenza delle serie per  $\varepsilon$  piccolo e i coefficienti della serie ammettono una semplice rappresentazione grafica.

(2) I coefficienti di Taylor in  $\varepsilon$  per  $\varepsilon \vec{\Xi}_\varepsilon, \varepsilon \vec{\Delta}_\varepsilon$  vengono espressi all’ordine  $n$  come somma di addendi che, *se stimati individualmente*, conducono ad una maggiorazione di  $\vec{\Xi}^{(n)}(\vec{\psi}), \vec{\Delta}^{(n)}(\vec{\psi})$  che cresce come una potenza di  $n!$ . Quindi è necessario verificare l’esistenza di cancellazioni prima di eseguire le stime di  $|\vec{\Xi}^{(n)}(\vec{\psi})|, |\vec{\Delta}^{(n)}(\vec{\psi})|$ .

(3) La rappresentazione grafica dei coefficienti  $\vec{\Xi}^{(n)}(\vec{\alpha}), \vec{\Delta}^{(n)}(\vec{\alpha})$  è conveniente perché consente di vedere quali siano i raggruppamenti dei termini, che ad essi contribuiscono, utili per migliorare la loro stima in  $B^n$ , con  $B < \infty$  e quindi ottenere la effettiva esistenza delle funzioni  $\vec{\Xi}_\varepsilon, \vec{\Delta}_\varepsilon$  per  $|\varepsilon| < B^{-1}$ .

Il metodo ora descritto segue le idee combinatorie del lavoro di Siegel, [2], integrandolo con i lavori successivi [10, 9, 11].

## 7 Serie di Lindstedt-Newcomb

La serie di Lindstedt-Newcomb, [6, Ch.IX], si ottiene imponendo che le Eq.(6.1) risolvano le equazioni del moto per la Hamiltoniana  $H_\varepsilon(\vec{A}, \vec{\alpha}) = \frac{1}{2J}\vec{A} \cdot \vec{A} + \varepsilon P(\vec{\alpha})$ : ossia sostituendo (con  $t = 0$ , il che basta) si trova

$$\begin{aligned} (\vec{\omega}_0 \cdot \vec{\partial}_{\vec{\psi}})\vec{\Delta}_\varepsilon(\vec{\psi}) &= J^{-1}\vec{\Xi}_\varepsilon(\vec{\psi}) \\ (\vec{\omega}_0 \cdot \vec{\partial}_{\vec{\psi}})\vec{\Xi}_\varepsilon(\vec{\psi}) &= -\vec{\partial}_{\vec{\alpha}}P(\vec{\psi} + \varepsilon\vec{\Delta}_\varepsilon(\vec{\psi})) \end{aligned} \quad (7.1)$$

ove  $\vec{\partial}_{\vec{\alpha}}P(\vec{\psi} + \varepsilon\vec{\Delta}_\varepsilon(\vec{\psi})) = \vec{\partial}_{\vec{\alpha}}P(\vec{\alpha})|_{\vec{\alpha}=\vec{\psi}+\varepsilon\vec{\Delta}_\varepsilon(\vec{\psi})}$ . Allora:

**Teorema 4:** (*Lindstedt-Newcomb*) *Esiste al più una soluzione  $\vec{\Delta}_\varepsilon(\vec{\psi}), \vec{\Xi}_\varepsilon(\vec{\psi})$  delle Eq.(7.1) che sia analitica in  $\varepsilon$  presso  $\varepsilon = 0$ , e con media nulla su  $\vec{\psi}$ . I coefficienti di Taylor a  $\varepsilon = 0$  sono comunque unicamente determinati, indipendentemente dall'effettiva esistenza o meno della soluzione. La serie formale di potenze in  $\varepsilon$  che risulta è detta "serie di Lindstedt".*

*Dim.:* Siano  $\vec{\Xi}^{(k)}(\vec{\psi}), \vec{\Delta}^{(k)}(\vec{\psi}), k \geq 1$ , i coefficienti del  $k$ -mo ordine dello sviluppo di Taylor di  $\varepsilon\vec{\Xi}_\varepsilon(\vec{\psi}), \varepsilon\vec{\Delta}_\varepsilon(\vec{\psi})$  in potenze di  $\varepsilon$ . Imponendo la Eq.(7.1), (e cioè che Eq.(6.1) risolva le equazioni del moto per  $H_\varepsilon$ ) si ottiene dalla prima  $\vec{\Xi}_\varepsilon(\vec{\psi}) = J(\vec{\omega}_0 \cdot \vec{\partial}_{\vec{\psi}})\vec{\Delta}_\varepsilon(\vec{\psi})$ ; combinando questa relazione con la seconda Eq.(7.1) segue la condizione necessaria e sufficiente

$$(\vec{\omega}_0 \cdot \vec{\partial}_{\vec{\psi}})^2\vec{\Delta}_\varepsilon(\vec{\psi}) = -J^{-1}\vec{\partial}_{\vec{\alpha}}P(\vec{\psi} + \varepsilon\vec{\Delta}_\varepsilon(\vec{\psi})), \quad (7.2)$$

Quindi la sola incognita è  $\vec{\Delta}_\varepsilon$ . Ponendo  $\varepsilon\vec{\Delta}_\varepsilon(\vec{\psi}) = \varepsilon\vec{\Delta}^{(1)}(\vec{\psi}) + \varepsilon^2\vec{\Delta}^{(2)}(\vec{\psi}) + \dots$  euguagliando i coefficienti dello sviluppo in  $\varepsilon$  dei due membri della Eq.(7.2) si ottengono relazioni ricorsive per i coefficienti  $\vec{\Delta}^{(k)}(\vec{\psi})$  (e quindi anche per  $\vec{\Xi}^{(k)}(\vec{\psi})$ ). Si trova per  $k = 1$

$$\vec{\Delta}^{(1)}(\vec{\psi}) = - \sum_{0 < |\vec{\nu}| < N} \frac{iJ^{-1}\vec{\nu}}{(i\vec{\omega}_0 \cdot \vec{\nu})^2} f_{\vec{\nu}} e^{i\vec{\nu} \cdot \vec{\psi}}. \quad (7.3)$$

quale unica soluzione a media nulla (si noti che la somma è finita perché  $P$  è un polinomio trigonometrico di grado  $< N$ , per ipotesi).

E in generale si trova il  $k$ -mo ordine sviluppando in potenze di  $\vec{\Delta}_\varepsilon(\vec{\psi})$  la funzione  $\vec{\partial}_{\vec{\alpha}}P(\vec{\psi} + \varepsilon\vec{\Delta}_\varepsilon(\vec{\psi}))$  e poi sviluppando  $\varepsilon\vec{\Delta}_\varepsilon(\vec{\psi})$  in potenze di  $\varepsilon$ . Si

ottiene, per  $k > 1$ ,

$$(\vec{\omega}_0 \cdot \vec{\partial}_{\vec{\psi}})^2 \vec{\Delta}^{(k)}(\vec{\psi}) = -J^{-1} \sum_{m=1}^{\infty} \frac{1}{m!} \vec{\partial}_{\vec{\alpha}}^{m+1} f(\vec{\psi}) \sum_{\substack{k_1, \dots, k_m \geq 1 \\ k_1 + \dots + k_m = k-1}} \prod_{r=1}^m \vec{\Delta}^{(k_r)}(\vec{\psi}). \quad (7.4)$$

Qui la notazione è semplificata per facilitare la lettura della formula. Se si intende calcolare la componente  $j$  di  $\vec{\Delta}^{(k)}(\vec{\psi})$  e se si pone  $\partial_j \stackrel{def}{=} \frac{\partial}{\partial \alpha_j}$  allora  $\vec{\partial}_{\vec{\alpha}}^{1+m} P(\vec{\psi})$  denota il tensore  $\partial_j \partial_{j_1} \dots \partial_{j_m} P(\vec{\alpha})$  calcolato in  $\vec{\alpha} = \vec{\psi}$  e gli indici  $j_1, \dots, j_m$  di questo tensore devono intendersi contratti con gli indici  $j_1, \dots, j_m$  che determinano le componenti dei vettori  $\vec{\Delta}^{(k_r)}(\vec{\psi})$ ; quindi se  $\partial_{j_1 \dots j_m} \stackrel{def}{=} \partial_{j_1} \dots \partial_{j_m}$  e  $k_r \geq 1$  si intende:

$$\begin{aligned} (\vec{\partial}_{\vec{\alpha}}^{m+1} f(\vec{\psi})) \sum_{k_1 + \dots + k_m = k-1, k_i > 0} \prod_{r=1}^m \vec{\Delta}^{(k_r)}(\vec{\psi})_{j_0} &= \\ &= \sum_{j_1, \dots, j_m=1}^{\ell} \sum_{k_1 + \dots + k_m = k-1, k_i > 0} \partial_{j j_1 \dots j_m} P(\vec{\psi}) \prod_{r=1}^m \Delta_{j_r}^{(k_r)}(\vec{\psi}), \end{aligned} \quad (7.5)$$

mentre, naturalmente, l'indice  $j$  resta non contratto e indica la componente di  $\vec{\Delta}^{(k)}(\vec{\psi})$  che si calcola.

Il membro di destra della Eq.(7.5) dipende solo dai coefficienti  $\vec{\Delta}^{(p)}(\vec{\psi})$  con  $p < k$ . Se quindi si chiama  $\vec{\Phi}^{(k)}(\vec{\psi})$  il secondo membro della Eq.(7.4) vediamo che, dato  $k > 1$ ,  $\vec{\Phi}^{(k)}(\vec{\psi})$  è determinato se sono noti i coefficienti  $\vec{\Phi}^{(p)}(\vec{\psi})$  con  $p < k$ . Quindi  $\vec{\Delta}^{(k)}(\vec{\psi})$  è determinato dall'equazione

$$(\vec{\omega}_0 \cdot \vec{\partial}_{\vec{\psi}})^2 \vec{\Delta}^{(k)}(\vec{\psi}) = \vec{\Phi}^{(k)}(\vec{\psi}) \quad (7.6)$$

che può essere risolta (e in un sol modo se si richiede che la soluzione abbia media nulla su  $\vec{\psi}$ ) come  $-\sum_{\vec{v} \neq \vec{0}} \frac{\vec{\Phi}^{(k)} e^{i\vec{v} \cdot \vec{\psi}}}{(\vec{\omega}_0 \cdot \vec{v})^2}$  se e solo se  $\vec{\Phi}^{(k)}(\vec{\psi})$  ha media nulla su  $\vec{\psi}$  (cioè se  $\vec{\Phi}_0^{(k)} = 0$ ) e in tal caso segue (induttivamente) dalla Eq.(7.4) che  $\vec{\Delta}^{(k)}(\vec{\psi}) = \sum_{0 < |\vec{v}| < kN} \vec{\Delta}_{\vec{v}}^{(k)} e^{i\vec{v} \cdot \vec{\psi}}$  è un polinomio trigonometrico di grado  $< kN$ .

Dunque rimane da dimostrare che la media di  $\vec{\Phi}^{(p)}(\vec{\psi})$  è nulla. Per questo basta osservare che  $P$  è per ipotesi pari in  $\vec{\psi}$  (ossia è un polinomio trigonometrico in soli coseni): quindi  $\vec{\Delta}^{(1)}(\vec{\psi})$  è per la Eq.(7.3) dispari in  $\vec{\psi}$ . Supponendo induttivamente che per  $p < k$  le funzioni  $\vec{\Delta}^{(p)}(\vec{\psi})$  siano dispari, dalla Eq.(7.5) si evince che  $\vec{\Phi}^{(k)}$  è combinazione lineare di prodotti di  $m$  funzioni dispari

per la derivata  $m + 1$ -ma di una funzione pari. Ne segue che  $\vec{\Phi}^{(k)}(\vec{\psi})$  è dispari e quindi ha media nulla su  $\vec{\psi}$ .

*Q.E.D.*

## 8 Grafici e serie di Linstedt

La moltitudine di indici nella Eq.(7.4) rende utile trovarne una rappresentazione sintetica. La costruzione è più facile se invece delle funzioni  $\vec{\Delta}^{(k)}(\vec{\psi})$  si studiano le loro trasformate di Fourier  $\vec{\Delta}_{\vec{v}}^{(k)}$ , perché l'inversione dell'operatore  $(\vec{\omega}_0 \cdot \vec{\partial}_{\vec{v}})^2$  è rappresentata dalla semplice divisione per  $-(\vec{\omega}_0 \cdot \vec{v})^2$ .

Si riscriva la Eq.(7.4)

$$\Delta_{j\vec{v}}^{(k)} = \frac{\delta_{jj_0}}{J(\vec{\omega}_0 \cdot \vec{v})^2} \sum_{m=1}^{\infty} \sum_{\vec{v}_0 + \vec{v}_1 + \dots + \vec{v}_m = \vec{v}} \frac{1}{m!} i\vec{\omega}_0^j f_{\vec{v}_0} \sum_{\substack{k_1, \dots, k_m \geq 1 \\ k_1 + \dots + k_m = k-1}} \prod_{s=1}^m i\vec{\omega}_0 \cdot \vec{\Delta}_{\vec{v}_s}^{(k_s)} \quad (8.1)$$

e si noti che può essere rappresentata via la Fig.1:

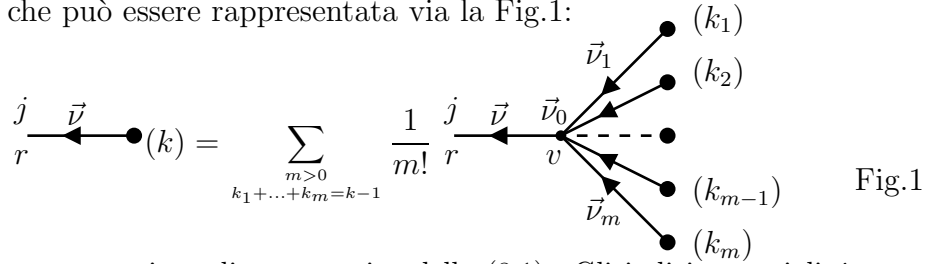


Fig.1

Fig.1: Rappresentazione diagrammatica della (8.1). Gli indici tensoriali  $j_s$  contratti non sono segnati.

ove il membro di sinistra è un simbolo per la componente  $\Delta_{\vec{v},j}^{(k)}$  di  $\vec{\Delta}_{\vec{v}}^{(k)}$  mentre a destra è rappresentato un “semplice albero” con

- (a) una “radice”  $r$  con indice “tensoriale”  $j$  (in  $1, 2, \dots, \ell$ ),
- (b) un “ramo orientato” verso la radice,  $\lambda_0 \equiv rv$ , che esce da un “nodo” (o “vertice”)  $v$  con “momento”  $\vec{v}_0$  e con un indice di tensore  $j_0$  (non segnato in Fig.1), e
- (c)  $m_v$  altri “rami” che “entrano in  $v$ ”, ciascuno dei quali ha un “momento”  $\vec{v}_s$ ,  $s = 1, \dots, m$  e possiede un indice di tensore rispettivo  $j_1, \dots, j_m$ ;
- (d) e i “vertici estremi” portano indicato il loro “ordine”  $k_s$ ,  $s = 1, \dots, m$ .
- (e) la lunghezza di rami e angoli fra rami adiacenti sono arbitrari.
- (f) Il nodo  $v$  simbolizza il tensore con elementi

$$\Phi_{v;j_0,j_1,\dots,j_m} = \frac{1}{m!} J^{-1}(i\vec{\omega}_0)_{j_0} f_{\vec{v}_0} \prod_{s=1}^m (i\vec{\omega}_0)_{j_s}, \quad (8.2)$$

(g) Il ramo uscente dal nodo  $v$  rappresenta il *propagatore*, ossia la matrice diagonale

$$\frac{\delta_{jj_0}}{(\vec{\omega}_0 \cdot \vec{v})^2} \quad (8.3)$$

(h) Il ramo uscente dal dischetto nel membro di sinistra rappresenta  $\vec{\Delta}_{\vec{v},j}^{(k)}$ ; e i rami che escono dai dischetti di destra con etichette  $(k_s)$ , c.f.r. Fig.1, rappresentano  $\vec{\Delta}_{j_s, \vec{v}_s}^{(k_s)}$ ,  $s = 1, \dots, mv$ .

(i) La relazione  $\vec{v} = \vec{v}_0 + \vec{v}_1 + \dots + \vec{v}_m$  verrà interpretata come “conservazione del momento” e si dirà che sul ramo  $\lambda_0$  fluisce una “corrente”  $\vec{v}$  uguale alla somma dei momenti del nodo  $v_0$  e delle correnti  $\vec{v}_i$  che emergono dai nodi estremi. Riflette, in ultima analisi, la simmetria per traslazione del toro.

I rami uscenti dagli  $m$  vertici estremi della Fig.1 possono essere interpretati graficamente allo stesso modo: osservando che il membro di destra della Eq.(8.1) è multilineare nelle  $\vec{\Delta}_{\vec{v}_s}^{k_s}$  ognuno dei rami entranti in  $v$  viene rappresentato come il membro di sinistra della Fig.1. Quindi iterando il procedimento si perviene, Fig.2, a rappresentare  $\Delta_{\vec{v},j}^{(k)}$  come risultato di una “somma su alberi” con rami orientati verso la radice e con  $k$  vertici  $v_i$ , ognuno dei quali ha un momento di nodo  $\vec{v}_{v_i}$ .

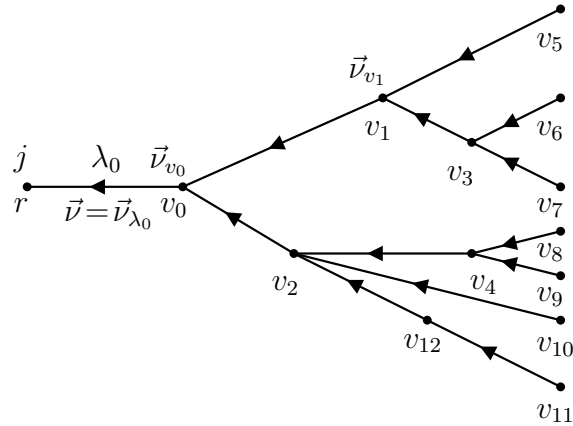


Fig.2

Fig.2: Un albero  $\vartheta$  con  $m_{v_0} = 2, m_{v_1} = 2, m_{v_2} = 3, m_{v_3} = 2, m_{v_4} = 2, m_{v_{12}} = 1$  e  $k = 13$ , a alcune decorazioni. Solo due momenti di nodo sono segnati esplicitamente (su  $v_0$  e  $v_1$ ) e sul ramo radice è segnata la “corrente” totale  $\vec{v} = \sum_v \vec{v}_n$ . Le frecce stabiliscono un ordine parziale,

Sui rami  $\lambda = v'v$  (con  $v'$  nodo immediatamente seguente  $v$ ) circola una “corrente” che è data dalla somma  $\vec{v}(\lambda) \stackrel{def}{=} \sum_{w \leq v} \vec{v}_w$  ove  $w \leq v$  significa che

il nodo  $w$  precede  $v$  nell'ordine parziale stabilito dalla orientazione dei rami. I vertici estremi  $v_f$  rappresentano  $\Delta_{\vec{v}_f, j_{v_f}}^{(1)}$ .

Se  $\lambda = v'v$  è un ramo allora la costruzione associa agli estremi di  $\lambda$  indici tensoriali  $j_{v'}$  e  $j_v$ . Tali indici sono indici tensoriali che si devono immaginare sommati, variando fra  $1, 2, \dots, \ell$ : a causa del propagatore “diagonale” questo significa porre  $j_{v'} = j_v$  e sommare su  $j_v = 1, 2, \dots, \ell$ . *Questo conduce ad una drammatica semplificazione* assegnando ad ogni albero un contributo al calcolo di  $\Delta_{\vec{v}, j}^{(k)}$  dato da  $k! \prod_v \frac{1}{m_v!}$  moltiplicato per

$$i \frac{1}{k!} \prod_{\lambda=v'v} \frac{\vec{v}_{v'} \cdot \vec{v}_v}{(\vec{\omega} \cdot \vec{v}(\lambda))^2} \prod_v J^{-1} f_{\vec{v}_v} \quad (8.4)$$

dove si immagina, per uniformità di notazione, che la radice  $r$  (che non è un nodo) abbia un momento  $\vec{v}_r = \vec{e}(j)$  uguale al versore della direzione  $j$ .

È opportuno immaginare che i rami che entrano in un nodo  $v$  siano contraddistinti da un indice variante da 1 a  $k$  che li distingue, e che due alberi che si possono sovrapporre facendo ruotare i rami attorno ai nodi nei quali entrano sono equivalenti. Allora il numero di modi di attribuire indici diversi ai rami è  $k! \prod_v \frac{1}{m_v!}$  e quindi se si immagina che gli alberi siano costruiti da rami distinguibili (ossia se si immagina che i  $k$  rami siano “presi da un deposito di rami” contraddistinti da un indice che li numera da 1 a  $k$ ) si può dire che

$$\Delta_{\vec{v}, j}^{(k)} = \frac{i}{k!} \sum_{\vartheta} \prod_{\lambda=v'v} \frac{\vec{v}_{v'} \cdot \vec{v}_v}{(\vec{\omega} \cdot \vec{v}(\lambda))^2} \prod_v J^{-1} f_{\vec{v}_v} \quad (8.5)$$

ove la somma verte su tutti gli alberi con rami distinti, coerentemente orientati e decorati dagli indici  $\vec{v}_v$  di momento: gli indici di componente sono d'ora in poi eliminati perché inclusi nei prodotti scalari  $\vec{v}_{v'} \cdot \vec{v}_v$ . Poiché per ipotesi  $f_{\vec{v}} = 0$  se  $|\vec{v}| \geq N$  si penserà sempre  $|\vec{v}_v| < N$ .

Ovviamente la somma in Eq.(8.5) verte su un gran numero di termini uguali: se immaginiamo di costruire gli alberi con radice e con  $k$  nodi e  $k$  linee ma senza distinguere fra i rami e poi di assegnare indici che distinguano i rami, otteniamo che il numero di alberi così costruiti e non sovrapponibili mediante rotazioni dei rami attorno ai loro nodi è  $k^{k-1}$ , esattamente: affermazione riducibile alla formula di Cayley che conta in  $k^{k-2}$  gli alberi (banalmente diversi dai presenti) con  $k-1$  rami e  $k$  nodi differenti uno dei quali è privilegiato.<sup>8</sup> Si è così ottenuto:

<sup>8</sup>Senza ricorrere alla formula di Cayley una stima del numero di alberi a nodi

**Teorema 5:** (degli alberi) Se  $\Theta(k, \vec{\nu}, j)$  è la famiglia di alberi  $\vartheta$  con  $k$  rami distinti e con i nodi decorati da momenti  $\vec{\nu}_v$ , con  $|\vec{\nu}_v| < N$  e  $\vec{\nu} = \sum_{v < r} \vec{\nu}_v$ ,  $|\vec{\nu}| < kN$  e momento  $\vec{e}_r$  sulla radice, e se si definisce il valore  $\text{Val}(\vartheta)$  di un albero la Eq(8.5) si ha

$$\Delta_{\vec{\nu}, j}^{(k)} = \sum_{\vartheta \in \Theta(k, \vec{\nu}, j)} \text{Val}(\vartheta) \quad (8.6)$$

e la somma verte su  $k^{k-1}$  alberi da decorare con le  $< (2N)^{\ell k}$  decorazioni possibili (ossia dai possibili  $\vec{\nu}_v \neq \vec{0}$ ).

Questo risultato è oggi noto come “serie di Lindstedt” sebbene sia solo una delle tante, si veda [12] per un’alternativa, possibili rappresentazioni dei coefficienti  $\Delta_{\vec{\nu}, j}^{(k)}$  la cui univoca definizione segue dal teorema 4.

## 9 Stima di Siegel. Piccoli denominatori.

Il teorema KAM si riduce, pertanto, alla stima dei valori degli alberi decorati con  $k$  nodi (e quindi  $k$  rami). Il numero di decorazioni è  $\leq (2N)^{\ell k}$  (perché le decorazioni sono gli indici  $\vec{\nu}_v$  di momento dei nodi e  $|\vec{\nu}| < N$  per ipotesi), e il numero di alberi non decorati (a rami distinti) è  $k^{k-1} < 4^k k!$  (c.f.r. nota<sup>8</sup>, i fattori  $\vec{\nu}_{v'} \cdot \vec{\nu}_v$  sono non superiori a  $N^2$  e i fattori  $J^{-1} f_{\vec{\nu}}$  sono non superiori a  $F = \max_{\vec{\nu}} J^{-1} |f_{\vec{\nu}}|$ ).

Pertanto la difficoltà per ottenere una stima  $|\Delta_{\vec{\nu}, j}^{(k)}| < B^k$  per qualche  $B$  sta nella presenza in Eq.(8.4) del  $\prod_{\lambda} \frac{1}{(\vec{\omega} \cdot \vec{\nu}(\lambda))^2}$ , che può essere molto grande (di ordine  $k!^a$  con  $a > 0$ ) anche tenendo conto della Eq.(2.3): questo costituisce la “difficoltà dei piccoli denominatori”, si veda l’esempio in Appendice A.

Occorre dunque collezionare famiglie di alberi in  $\Theta(k, \vec{\nu}, j)$  e sommarne i valori prima di tentare una loro stima.

L’idea è suggerita dal risultato di Siegel che nel nostro contesto può essere pensato così.

Definiamo la “proprietà  $\mathcal{S}$ ”: un albero decorato  $\vartheta$  possiede la proprietà  $\mathcal{S}$  se dati due rami  $\lambda < \lambda'$  (ossia con  $\lambda$  che precede  $\lambda'$  nell’ordine parziale dell’albero) si ha

$$\vec{\nu}(\lambda') \neq \vec{\nu}(\lambda), \quad \text{“proprietà di Siegel”} \quad (9.1)$$

---

distinti è  $k!4^k$ , peggiore di un fattore  $\sim (4e^{-1})^k k^{\frac{3}{2}} \sqrt{2\pi}$ .

allora

**Teorema 6:** (di Siegel-Pöschel) Si consideri il contributo  $\Delta_{\vec{v},j}^{(k;\mathcal{S})}$  a  $\Delta_{\vec{v},j}^{(k)}$  dei soli alberi con la proprietà  $\mathcal{S}$ : si avrà

$$\sum_{\vec{v} \neq \vec{0}} |\Delta_{\vec{v},j}^{(k;\mathcal{S})}| < B_0^k \quad (9.2)$$

con  $B_0 = (2N)^\ell \ell N^2 2^2 C^2 F \prod_{n=0}^{-\infty} 2^{-2(n-1)2^{\frac{n}{\tau}}\gamma}$ , e  $\gamma = 2 \cdot 2^{\frac{1}{\tau}} N$ .

*Dim.* (da [10, 11]) Dato un albero decorato  $\vartheta \in \Theta(k, \vec{v}, j)$ , si dirà che una corrente  $\vec{v}$  in un ramo di  $\vartheta$  è  $n$ -singolare se  $2^{n-1} < C|\omega_0 \cdot \vec{v}| \leq 2^n$ ,  $n = 0, -1, \dots$  e per  $n = 1$  se  $1 < C|\omega_0 \cdot \vec{v}|$ . Sia  $N(k, n)$  il numero di rami  $n$ -singolari.

Il punto chiave è che tutti i momenti di nodo devono avere  $|\vec{v}_v| < N$ . Quindi perché in un ramo  $\lambda = v'v$  circoli corrente  $\vec{v}(\lambda)$  che sia  $n$ -singolare è necessario che il nodo  $v$  sia preceduto da *almeno*  $N^{-1}2^{-\frac{n}{\tau}}$  nodi. Una volta che un ramo  $n$ -singolare è stato generato sono necessari circa altrettanti nodi per generarne un altro (per le proprietà  $\mathcal{S}$  e Diofantea) e così via. Poiché il numero totale di nodi è  $k$  ne consegue che il numero di rami  $n$ -singolari è stimabile proporzionalmente a  $kN2^{\frac{n}{\tau}}$  e per induzione si può verificare che la costante di proporzionalità può essere scelta  $2^{1+\frac{1}{\tau}}$ .

Invero supponiamo che  $N(k, n) \leq 2^{1+\frac{1}{\tau}} N k 2^{\frac{n}{\tau}} - 1$  se  $N(k, n) > 0$ . Ma se  $N(k, n) > 0$  la corrente  $\vec{v}$  dei rami singolari deve essere  $Nk \geq |\vec{v}| > 2^{-\frac{n}{\tau}}$  per la Eq.(2.3) e quindi l'affermazione è vera se  $\vartheta$  contiene un solo ramo  $n$ -singolare (e in questo caso si potrebbe fare a meno del fattore  $2^{\frac{1}{\tau}}$ ).

Se  $\vartheta$  contiene più di un ramo  $n$ -singolare deve avere più di un nodo: siano  $\vartheta_1, \vartheta_2, \dots, \vartheta_s$  gli  $s \geq 1$  sottoalberi che convergono nel nodo  $v_0$  che precede la radice di  $\vartheta$ .

Due possibilità: il ramo della radice è  $n$ -singolare o no. Se no allora il risultato segue per ipotesi induttiva (su  $k$  essendo vero per  $k = 1$ ): si noti il ruolo chiave che è tenuto dal  $-1$  che appare nell'ipotesi induttiva.

Oppure il ramo  $\lambda_0$  della radice è  $n$ -singolare. In questo caso o almeno due degli alberi  $\vartheta_1, \vartheta_2, \dots, \vartheta_s$  contengono un ramo singolare e allora il risultato è conseguenza dell'ipotesi induttiva (grazie ancora al  $-1$  nell'ipotesi induttiva) oppure uno solo, ad esempio  $\vartheta_1$ .

In tale ultimo caso si risalga  $\vartheta_1$  fino ad incontrare un ramo  $n$ -singolare  $\bar{\lambda}$  di corrente  $\vec{v}(\bar{\lambda})$ . Si consideri l'insieme  $\tilde{\vartheta}$  ottenuto cancellando  $\bar{\lambda}$  e tutti i rami che lo precedono. La somma di tutti i momenti di nodo di  $\tilde{\vartheta}$  deve



essere  $\vec{v} - \vec{v}(\bar{\lambda})$ . Poiché  $\lambda_0, \bar{\lambda}$  sono  $n$  singolari ma  $\vec{v} \neq \vec{v}(\bar{\lambda})$  si avrà  $0 < C|\vec{\omega}_0 \cdot \vec{v} - \vec{\omega}_0 \cdot \vec{v}(\bar{\lambda})| < 2 \cdot 2^n$ .

Ne consegue che  $\tilde{\vartheta}$  deve contenere almeno  $k_1 = \frac{1}{N} 2^{-\frac{n+1}{\tau}}$  nodi (e rami). Quindi i rami cancellati sono al più  $k - k_1$  rami e per l'ipotesi induttiva al più  $2^{1+\frac{1}{\tau}} N 2^{\frac{n}{\tau}} (k - k_1) - 1$  rami  $n$ -singolari, ossia al più

$$2^{1+\frac{1}{\tau}} N 2^{\frac{n}{\tau}} \left(k - \frac{1}{N} 2^{-\frac{n+1}{\tau}}\right) - 1 \leq 2^{1+\frac{1}{\tau}} N 2^{\frac{n}{\tau}} k - 2 \quad (9.3)$$

aggiungendo 1 a questo numero, per contare la linea  $n$ -singolare  $\lambda_0$ , si ottiene il risultato.

Il teorema 6 ora segue:  $\Delta_{\vec{v},j}^{(k;S)}$  si stima valutando il valore di un singolo albero con

$$\frac{1}{k!} (2N)^{\ell k} N^{2k} (C^2 F)^k \prod_{n=0}^{-\infty} 2^{-2(n-1) \cdot N(k,n)} \quad (9.4)$$

Stimando il numero di alberi non decorati con  $4^k k!$  e usando  $N(k, n) \leq 2Nk 2^{\frac{1}{\tau}} 2^{\frac{n}{\tau}}$  e i prodotti scalari  $|\vec{v}_v \cdot \vec{v}_v| \leq \ell N^2$  si trova, se  $\gamma \stackrel{def}{=} 2^{1+\frac{1}{\tau}} N$ ,

$$\sum_{\vec{v} \neq \vec{0}} |\Delta_{\vec{v},j}^{k,S}| \leq 4^k (2N)^{\ell k} (\ell N^2)^k (C^2 F)^k \left( \prod_{n=0}^{-\infty} 2^{-2(n-1) 2^{\frac{n}{\tau}} \gamma} \right)^k \quad (9.5)$$

Si noti che la somma sulle decorazioni degli alberi implica la somma sui momenti dei nodi (con la condizione che  $|nN(\lambda)| \neq \vec{0}$  for all  $\lambda$ ) e quindi anche la somma su  $\vec{v}$ .

*Q.E.D.*

**Teorema 7:** *La proprietà  $\mathcal{S}$ , Eq.(9.1), può essere indebolita ammettendo la possibilità dell'esistenza di coppie di rami  $\lambda < \lambda'$  con uguale corrente  $\vec{v}(\lambda) = \vec{v}(\lambda')$  se l'albero  $\tilde{\vartheta}$  ottenuto "potando" tutti i rami  $\leq \lambda$  ha un numero di nodi  $k_1 \geq 2^{-\frac{n+c}{\tau}} N^{-1}$  con  $c > 0$  intero arbitrario. In tal caso il numero di rami di scala  $n$  è  $N(k, n) \leq 2^{1+\frac{c}{\tau}} Nk 2^{\frac{n}{\tau}}$ .*

*Osservazione (1)* Si potrebbe fissare  $c = 1$ , ma la scelta che verrà fatta nel §10 dovrà essere  $c > 1$ , ad esempio  $c = 2$ .

(2) La dimostrazione è la stessa del teorema 6 perché la Eq.(9.3) è stata derivata inferendo che  $\tilde{\vartheta}$  deve contenere almeno  $k_1$  rami.

## 10 Analisi multiscala delle singolarità

I teoremi 6 e 7 mostrano che il vero problema sta nello stimare i contributi a  $\Delta_{\vec{v},j}^{(k)}$  dovuto agli alberi che contengono rami di ugual corrente che si seguono nell'ordine parziale dell'albero e che hanno pochi rami intermedi.

Già lo studio della stima di Siegel mostra che il problema è un problema di *analisi multiscala*. Si dividono i rami in classi di scala  $1, 0, -1, \dots$ : i rami  $\lambda$  in cui fluisce una corrente  $\vec{v}(\lambda)$  tale che

$$2^{n-1} < C|\vec{\omega}_0 \cdot \vec{v}(\lambda)| \leq 2^n, \quad n \leq 0 \quad (10.1)$$

hanno scala  $n \leq 0$  mentre se  $1 < C|\vec{\omega}_0 \cdot \vec{v}(\lambda)|$  si dice che  $\lambda$  ha scala 1. In tal modo si viene ad aggiungere ad ogni ramo degli alberi un *indice di scala*.

Convieni ora immaginare di racchiudere in un contorno chiuso un insieme connesso di rami (e quindi dei nodi loro estremi) che hanno scala  $\geq n$  fra i quali almeno uno di essi ha scala esattamente  $n$  (nella nostra convenzione  $n$  è solitamente negativo) e che non può essere ingrandito includendo altri rami di scala  $\geq n$ .

Chiameremo *bolle* l'insieme di rami (e loro estremi) racchiusi da tali contorni e un albero decorato è visualizzato con i rami racchiusi in contorni disposti gerarchicamente; guardando dentro un contorno si vedono solo rami di scala  $\geq n$  e almeno uno di scala  $n$  e nessuno dei rami che entrano nella bolla ha scala  $\geq n$ .

Si noti che in ogni bolla possono entrare più rami ma ne esce necessariamente uno solo o nessuno (perché le bolle sono disegnate su un albero).

La necessità di raccogliere i contributi di vari alberi prima di tentare una stima di  $\Delta_{\vec{v},j}^{(k)}$  è chiara se si intende utilizzare la stima di Siegel, opportunamente adattata. Convieni definire:

**Definizione:** una bolla  $S$  è detta “singolare su scala  $n$ ” se

(a) ha un solo ramo entrante  $\lambda_{in}$  la cui corrente ha scala  $n$  ed è uguale a quella del ramo uscente  $\lambda_{us}$ :  $\vec{v}(\lambda_{in}) = \vec{v}(\lambda_{us})$ ;

(b) il numero di rami contenuti nella bolla  $S$  è “piccolo” nel senso che il loro numero è  $\leq 2^{-\frac{c+n}{\tau}} N^{-1}$  con  $c \geq 2$  intero.

Osservazioni: (1) Necessariamente la scala  $n_S$  di una bolla singolare è  $n_S > n = n_{\lambda_{in}}$ , ossia i rami entranti nella bolla singolare  $S$  sono “più singolari” dei rami in  $S$ .

(2) Si noti che il piccolo numero di rami in una bolla  $S$  con singolarità di scala

$n$  fa sí che  $n_S \geq n + c$  perché nessun suo ramo  $\bar{\lambda}$  può avere corrente  $\vec{v}$  somma di una corrente  $|\vec{v}^0| \leq 2^{-\frac{n+c}{\tau}}$  e eventualmente una di scala  $n$  (essendo una corrente in una bolla singolare somma di al più  $2^{-\frac{n+c}{\tau}} N^{-1}$  momenti di nodo interni alla bolla, che generano un contributo alla corrente con  $|\vec{\omega}_0 \cdot \vec{v}^0| \geq 2^{\frac{n+c}{\tau}}$ ) e quindi  $C|\vec{\omega}_0 \cdot \vec{v}| \geq 2^{n+c} - 2^n$  che è  $\geq 2^{n+c-1}$ , ossia  $n_S \geq n + c$ .

(3) Gli alberi con bolle singolari non possono essere studiati (direttamente) con il metodo di Siegel.

Dato un albero decorato  $\vartheta_0$  consideriamo le sue bolle singolari. Si immagini di staccare il ramo  $\lambda_{S,in}$  entrante in una bolla singolare  $S$  e riattaccarlo ad un altro dei nodi interni alla bolla  $S$  *ma esterni* a altre bolle singolari eventualmente presenti all'interno di  $S$ . E si esegua questa operazione per ogni bolla singolare  $S$  di  $\vartheta_0$ : si ottiene una collezione di alberi alla quale aggiungiamo gli alberi che si possono ottenere invertendo simultaneamente i segni di tutti i momenti di nodo dei nodi interni ad una bolla singolare ma esterni alle bolle singolari eventualmente contenute in  $S$ . Chiameremo  $\mathcal{F}(\vartheta_0)$  la famiglia di alberi sí definita; e certamente se  $\vartheta \in \mathcal{F}(\vartheta_0)$  si ha  $\mathcal{F}(\vartheta) = \mathcal{F}(\vartheta_0)$ .

Gli indici che identificano i rami non sono toccati da queste operazioni e quindi resta una corrispondenza biunivoca fra i rami di uguale indice numerico degli alberi in  $\mathcal{F}(\vartheta_0)$

Se  $S$  è una bolla singolare le correnti dei rami interni a  $S$  possono cambiare al variare di  $\vartheta$  in  $\mathcal{F}(\vartheta_0)$ . Se  $\vec{v}(\lambda_{in})$  è il momento entrante in  $S$  e  $\bar{\lambda}$  è un ramo interno a  $S$  la sua scala inizialmente  $\bar{n}$  potrebbe cambiare perché il piccolo divisore per  $\bar{\lambda}$  può venire alterato da  $\vec{\omega}_0 \cdot \vec{v}(\bar{\lambda})$  in  $\vartheta \in \mathcal{F}(\vartheta_0)$  in  $\vec{\omega}_0 \cdot \vec{v}(\bar{\lambda}) \pm \vec{\omega}_0 \cdot \vec{v}_\lambda$  in un altro albero  $\vartheta' \in \mathcal{F}(\vartheta_0)$ .

Ma in molti casi questo *non accade*. Ad esempio se  $\vec{\omega}_0$  ha la proprietà per tutti gli interi  $n \leq 0$ , che chiamiamo *Diofantea forte*:

$$\max_{0 \geq p \geq n} |C\vec{\omega}_0 \cdot \vec{v} - 2^p| > 2^{n+1}, \quad \forall 0 < |\vec{v}| \leq 2^{-\frac{n+c}{\tau}}, \quad c \geq 2 \text{ intero} \quad (10.2)$$

allora, visto che  $S$  è  $n$ -singolare per la osservazione (2) precedente ogni suo ramo  $\bar{\lambda}$  ha un propagatore, se  $\sigma = 0, 1$ , con denominatore  $\geq 2^{c+n} - \sigma 2^n > 2^{n+c-1} > 2^{n+1}$  e quindi (per la Eq.(10.2)) il ramo  $\bar{\lambda}$  non cambia scala, pur cambiando la sua corrente. Si noti infatti che il significato della Eq.(10.2) è semplicemente la proprietà che  $C\vec{\omega}_0 \cdot \vec{v}$  cade in tutti gli intervalli  $[2^p, 2^{p-1}]$  a distanza  $> 2^n$  dagli estremi (per i  $\vec{v}$  rilevanti,  $|\vec{v}| \leq 2^{-\frac{n+c}{\tau}}$ ).

Supporremo, per semplicità valida la Eq.(10.2) che *insieme alla* Eq.(2.3) *con*  $\tau > \ell$  *è comunque una proprietà vera per tutti i vettori di rotazione*  $\vec{\omega}_0$

con la sola eccezione di un insieme di vettori  $\vec{\omega}_0 \in R^\ell$  di misura nulla, si veda §12 per il caso generale.

In questo caso, in definitiva, le scale di tutti i rami corrispondenti degli alberi in  $\mathcal{F}(\vartheta_0)$  non cambiano.

Indicando con  $S$  le bolle singolari sia  $M_S$  il numero di nodi contenuti in  $S$  ma non in eventuali bolle singolari interne a  $S$ . La famiglia  $\mathcal{F}(\vartheta)$  contiene  $\prod_S (2M_S) \leq e^{\sum_S 2M_S} \leq e^{2k}$  alberi diversi.

La presenza di bolle singolari modifica (sostanzialmente) la Eq.(9.4) in

$$\frac{1}{k!} e^{2k} 4^k (2N)^{\ell k} (\ell N^2)^k (C^2 F)^k \prod_{n=0}^{-\infty} 2^{-2(n-1) \cdot N_n} \quad (10.3)$$

ove  $N_n$  è il numero di rami di scala  $n$  (per il quale ora non si può usare la stima di Siegel, almeno non direttamente).

Si osservi che se  $S$  è una bolla  $n$ -singolare i cui rami interni hanno scala massima  $n_S$  e nella quale entra una corrente  $\vec{v}(\lambda_{S,in})$  di scala  $n_{\lambda_{S,in}} = n$  (ed è dunque  $n \leq n_S - c$  per la precedente osservazione (2)) i propagatori dei rami  $\bar{\lambda}$  interni a  $S$  hanno la forma  $(\vec{\omega}_0 \cdot \vec{v}_{\bar{\lambda}}^0 + \sigma \varepsilon_S)^{-2}$  con  $\varepsilon_S = \vec{\omega}_0 \cdot \vec{v}_{\lambda_{S,in}}$ ,  $\vec{v}_{\bar{\lambda}}^0$  uguale alla corrente che fluirebbe in  $\bar{\lambda}$  se la corrente immessa nella bolla dal ramo entrante fosse  $= \vec{0}$  (ossia se il ramo avesse scala  $-\infty$  in luogo di  $n$ ), e  $\sigma = 0, 1$  a seconda che  $\bar{\lambda}$  si trovi sul precorso che, sull'albero, porta dal ramo entrante a quello uscente da  $S$ .

Infine conviene ordinare gerarchicamente le bolle singolari in bolle di livello  $1, 2, \dots$ . Una bolla singolare è di livello 1 se non contiene alcuna bolla singolare; è di livello  $p > 1$  se contiene solo bolle di livello  $\leq p - 1$  e una almeno di tal livello. Siano  $\mathcal{L}_1, \mathcal{L}_2, \dots, \mathcal{L}_q$  le famiglie di bolle singolari dei vari livelli in  $\mathcal{F}(\vartheta_0)$ .

## 11 Cancellazioni

La stima Eq.(10.3) vale ancora se in tutti i propagatori *interni* a  $S$  la  $\varepsilon_S = \vec{\omega}_0 \cdot \vec{v}_{\lambda_{S,in}}$  viene sostituita da una variabile complessa  $z_S$  di modulo  $|z_S| \leq 2^{n_{\lambda_{S,in}}}$ . Riguardando la somma degli alberi in  $\mathcal{F}(\vartheta)$  come funzione  $\Phi$  di  $z_S$  si vede che  $\Phi$  si annulla per  $z_S = 0$  al secondo ordine.

È questa la *cancellazione* che permette buone stime. Avviene perché i valori degli alberi di  $\mathcal{F}(\vartheta_0)$  che si ottengono facendo variare il nodo  $v$  cui si attacca il ramo  $\lambda_{S,in} = vv_-$  che entra in  $S$  differiscono sia per il fattore  $\vec{v}_v \cdot \vec{v}_{v_-}$

sia perché i propagatori interni alla bolla  $S$  possono cambiare di  $\pm z_S$ . Ma la somma su  $v$  di  $\vec{v}_v \cdot \vec{v}_{v_-}$  è nulla perché  $\sum_v \vec{v}_v = \vec{0}$  (i vertici  $v$  sono tutti i vertici interni alla bolla e esterni alle eventuali bolle interne e quindi la somma in questione è nulla essendo uguale alla differenza  $\vec{v}_{\lambda_{in}} - \vec{v}_{\lambda_{us}} = \vec{0}$ ).

Quindi la funzione  $\Phi$  di  $z_S$  ha uno zero in  $z_S = 0$ : in realtà lo zero è doppio perché in  $\mathcal{F}(\vartheta_0)$  si considerano anche gli alberi ottenuti invertendo simultaneamente e indipendentemente i momenti dei nodi interni alle bolle singolari e quindi, tenendo conto che tutti i propagatori sono funzioni pari delle correnti, la  $\Phi$  è pari in  $z_S$ .

Quest'osservazione implica che la stima Eq.(10.3) è migliorata da un fattore  $2^{-2(n_S - n_{\lambda_S, in} - 1)}$ : perché la funzione  $\Phi$  viene calcolata per un valore  $|z_S| \leq 2^{n_{\lambda_S, in}}$  ma ha un raggio di convergenza almeno  $2^{n_S - 1}$  (perché questa è la grandezza minima delle quantità  $C|\vec{\omega}_0 \cdot \vec{v}^0(\bar{\lambda})|$  per rami di scala  $n_S$ ) e uno zero di secondo ordine a  $z_S = 0$  (quindi questo segue dal principio di massimo). Sommando anche sugli indici che distinguono i rami (al più  $k!4^k$ ) e le decorazioni  $\vec{v}_w$  sui nodi  $w$  di  $\vartheta$  si trova la stima

$$|\Delta_{\vec{v}, j}^{(k)}| \leq 4^k e^{2k} (2N)^{\ell k} N^{2k} (C^2 F)^k \left( 2^{-2(n_S - n_{\lambda_S, in} - 1)} \right) \left( \prod_{n=0}^{-\infty} 2^{-2(n-1) \cdot N_n} \right) \quad (11.1)$$

ove  $N_n$  è il numero di rami di scala  $n$ .

Vogliamo mostrare che il fattore dipendente da  $S$  in Eq(11.1) può essere migliorato in  $\prod_S \left( 2^{-2(n_S - n_{\lambda_S, in} - 1)} \right)$ , cioè vogliamo mostrare che le cancellazioni sono indipendenti.

Dato un albero  $\vartheta_0$  immaginiamo di eseguire la somma del  $\text{Val}(\vartheta)$  al variare di  $\vartheta \in \mathcal{F}(\vartheta_0)$  sommando prima sui  $\vartheta \in \mathcal{F}(\vartheta_0)$  al variare di  $\vartheta$  fra gli alberi in cui si modifica solo il contenuto delle bolle singolari di livello 1 (cioè si cambia solo il nodo cui si attacca il ramo entrante mantenendo tutte le decorazioni fisse). Allora, posto  $n_{\lambda_S} \stackrel{def}{=} n_{\lambda_S, in} = n_{\lambda_S, us}$ , si ha la stima

$$\frac{1}{k!} \left( \prod_{S \in \mathcal{L}_1} (2M_S) \right) N^{2k} (C^2 F)^k \prod_{S \in \mathcal{L}_1} \left( 2^{-2(n_S - n_{\lambda_S} - 1)} \right) \left( \prod_{n=0}^{-\infty} 2^{-2(n-1) \cdot N_n} \right) \quad (11.2)$$

perché le somme sui contenuti delle bolle singolari di  $\mathcal{L}_1$  sono indipendenti e quindi si ha uno zero per ogni  $z_S$ : ossia per quanto detto possiamo migliorare il fattore dipendente da  $S$  nella Eq.(11.1) in  $\prod_{S \in \mathcal{L}_1} \left( 2^{-2(n_S - n_{\lambda_S} - 1)} \right)$ .

Sommiamo poi  $\text{Val}(\vartheta)$  al variare di  $\vartheta \in \mathcal{F}(\tau_0)$  sugli alberi ottenuti modificando anche i contenuti delle bolle singolari in  $\mathcal{L}_2$ : ogni bolla singolare  $S$  di

secondo livello ha rami interni che sono funzioni analitiche di  $z_S = \vec{\omega}_0 \cdot \vec{\nu}(\lambda_{S,in})$  con raggio di convergenza  $2^{n_S-1}$  e che ammettono la stima Eq.(11.2) perché somme di valori di bolle in  $\mathcal{L}_1$  con rami entranti di scala  $n_i$  dunque con  $|z_i| < 2^{n_i}$ . La cancellazione implica che la Eq.(11.2) è migliorata in

$$\frac{1}{k!} \left( \prod_{S \in (\mathcal{L}_1 \cup \mathcal{L}_2)} (2M_S) \right) N^{2k} (C^2 F)^k \left( \prod_{S \in (\mathcal{L}_1 \cup \mathcal{L}_2)} 2^{-2(n_S - n_{\lambda_S} - 1)} \right) \left( \prod_{n=0}^{-\infty} 2^{-2(n-1) \cdot N_n} \right) \quad (11.3)$$

e così via. Si trova, sommando anche sugli indici che distinguono i rami e sui momenti dei nodi:

$$|\Delta_{\vec{\nu}_j}^{(k)}| \leq 4^k e^{2k} (2N)^{\ell k} N^{2k} (C^2 F)^k \left( \prod_S 2^{-2(n_S - n_{\lambda_S} - 1)} \right) \left( \prod_{n=0}^{-\infty} 2^{-2(n-1) \cdot N_n} \right) \quad (11.4)$$

Resta da vedere se il “peggioramento” sul valore di  $N_n$  dovuto alla presenza delle bolle singolari, è compensato dai fattori  $2^{-2(n_S - n_{\lambda_S} - 1)}$  dovuti alle cancellazioni.

Le bolle  $n$ -singolari sono sempre contenute in bolle (singolari o no) di scala  $n$ : se  $T$  è una bolla di scala  $n$  denotiamo  $m_T$  il numero di bolle  $n$ -singolari contenute in  $T$ .

Allora il numero  $N_n$  di rami di scala  $n$  si migliora via la stima di Siegel. Tale stima valuta  $2^{1+\frac{\epsilon}{\tau}} N 2^{-\frac{n}{\tau}} k \stackrel{def}{=} \overline{N}_n$  il numero di bolle di scala  $n$  contenute in un albero in cui non ci sono bolle  $n$ -singolari, *c.f.r.* §10, dedotta nel teorema 7, e si trova

$$N_n \leq 2^{2^{1+\frac{\epsilon}{\tau}}} N 2^{-\frac{n}{\tau}} k + \sum_{T, n_T=n} (-1 + m_T) \quad (11.5)$$

Invero la stima di Siegel (teoremi 6,7 del §9) implica  $N_n \leq \overline{N}_n + \sum_{T, n_T=n} m_T$ : basta immaginare di eliminare dall'albero tutte le bolle singolari di scala  $n$  e i rami contenuti (che hanno scala  $> n$ ) congiungendo in un unico ramo quello entrante e quello uscente da esse; si ottiene così un albero cui si può applicare il teorema 7.

Inoltre il numero di bolle di scala  $n$  (singolari o no) in un albero a  $k$  rami è stimato da  $\overline{N}_n(k) - 1$ , se esistenti; infatti se  $\overline{N}_n(k)$  è il numero di bolle (singolari o no) di scala  $n$  l'affermazione è vera per  $k = 1$ . Per induzione su  $k$ : o la radice dell'albero non è in una bolla di scala  $n$  e allora l'affermazione è vera per ipotesi induttiva; oppure è in una tale bolla e se essa è preceduta da almeno due bolle di scala  $n$  l'affermazione resta vera per l'ipotesi induttiva;

infine se la radice è in una bolla di scala  $n$  e il sottoalbero la cui radice entra nella bolla ne contiene una sola di scala  $n$  allora il numero di nodi nel sottoalbero è  $k_1 \leq k - 2^{-\frac{n+c}{\tau}} N$  che é sufficiente a ottenere la validità dell'ipotesi induttiva anche in questo caso.

Quindi  $N_n \leq 2\bar{N}_n + \sum_{T, n_T=n} (-1 + m_T)$ .

Si osservi ora che  $2 \sum_T (-1 + m_T) + \sum_{S \text{ singolare}} 2(n_{\lambda_S} - n_S) \leq 0$  e quindi

$$\sum_{\vec{\nu} \neq \vec{0}} |\Delta_{\vec{\nu},j}^{(k)}| \leq 4^k e^{2k} (2N)^{\ell k} N^{2k} (C^2 F)^k 2^{2k} \left( \prod_{n=0}^{-\infty} 2^{-2(n-1)2^{2+\frac{c}{\tau}} N 2^{-\frac{n}{\tau}}} \right)^k \quad (11.6)$$

e quindi la serie per  $\vec{\Delta}_\varepsilon$ , e addirittura per  $\sum_{\vec{\nu} \neq \vec{0}} \sum_{k=1}^{\infty} |\varepsilon|^k |\vec{\Delta}_{\vec{\nu},j}^{(k)}|$ , converge per

$$|\varepsilon| < 4e^2 (2N)^\ell N^2 (C^2 F) 2^2 \left( \prod_{n=0}^{-\infty} 2^{-2n 2^{2+\frac{c}{\tau}} N 2^{-\frac{n}{\tau}}} \right) \stackrel{def}{=} \varepsilon_{kam}^0 \quad (11.7)$$

ove  $c$  può essere scelto  $c = 2$ , ad esempio.

Si osservi, infine, che per costruzione  $\Delta_{\vec{\nu},j}^{(k)} \equiv 0$  se  $|\vec{\nu}| \geq kN$ , per l'ipotesi che la perturbazione  $f$  sia un polinomio trigonometrico di grado  $N$ . Ne segue, se  $\vec{m} = (m_1, \dots, m_\ell)$  e con notazioni naturali, che la serie di Fourier

$$\varepsilon \vec{\partial}_{\vec{\psi}}^{\vec{m}} \vec{\Delta}_\varepsilon(\vec{\psi}) = \sum_{k=1}^{\infty} \sum_{\vec{\nu}} \varepsilon^k (i\vec{\nu})^{\vec{m}} \vec{\Delta}_{\vec{\nu}}^{(k)} e^{i\vec{\nu} \cdot \vec{\psi}} \quad (11.8)$$

essendo maggiorata da  $\sum_{k=1}^{\infty} |\varepsilon|^k (kN)^{|\vec{m}|} \sum_{\vec{\nu} \neq \vec{0}} |\vec{\Delta}_{\vec{\nu}}^{(k)}|$ , converge uniformemente per  $|\varepsilon| < \varepsilon_{kam}^0$ . E, posto secondo la Eq.(7.1),  $\vec{\Xi}(\vec{\psi}) \stackrel{def}{=} J\vec{\omega}_0 \cdot \vec{\partial}_{\vec{\psi}} \vec{\Delta}(\vec{\psi})$ , questo dimostra il teorema KAM per l'Hamiltoniana Eq.(6.2):

**Teorema 8:** (*KAM per rotatori*) Se  $\vec{\omega}_0$  è una rotazione Diofantea verificante la Eq.(10.2) e se  $H_\varepsilon(\vec{A}, \vec{\alpha}) = \frac{1}{2J} \vec{A} \cdot \vec{A} + \varepsilon P(\vec{\alpha})$  con  $P$  polinomio trigonometrico pari di grado  $< N$  la serie di Lindstedt per il moto quasi periodico con spettro  $\vec{\omega}_0$  converge per  $\varepsilon$  abbastanza piccolo.

*Q.E.D.*

## 12 Osservazioni

(1) L'ipotesi *Diofantea forte*, Eq.(10.2), si elimina sfruttando l'evidente arbitrarietà del fattore 2 usato nella definizione delle scale: un qualsiasi  $\gamma > 1$  in luogo di  $\gamma = 2$  permetterebbe di pervenire alla stessa conclusione con sole

modifiche nelle costanti nelle stime conclusive Eq.(11.4). Anzi più in generale basta supporre che oltre alla proprietà Diofantea, Eq.(2.3), esista una successione  $\gamma_0, \gamma_{-1}, \dots$  con  $\gamma_p \in (2^{p-1}, 2^p]$  tale che:

$$|C\vec{\omega}_0 \cdot \vec{\nu} - \gamma_p| \geq 2^{n+1}, \quad \text{se } n \leq p \leq 0 \quad \text{e } 0 < |\vec{\nu}| \leq 2^{-\frac{n-3}{\tau}} \quad (12.1)$$

per tutti gli  $n \leq 0$  (e anche qui le potenze di 2 possono essere sostituite da potenze di un arbitrario numero  $> 1$  modificando varie costanti). Si dimostra, [13], che la proprietà che conduce alla Eq.(12.1) è sempre valida come conseguenza della Eq.(2.3) stessa.

(2) L'ipotesi  $h = \frac{1}{2J}\vec{A} \cdot \vec{A}$  può essere indebolita in  $\frac{M}{2}\vec{A} \cdot \vec{A} + \vec{\omega} \cdot \vec{A}$  con  $M$  matrice simmetrica (anche se singolare: *teorema KAM senza torsione*) se  $\vec{\omega}_0 = M\vec{A}_0 + \vec{\omega}$  per qualche  $\vec{A}_0$  e  $\vec{\omega}_0$  è Diofanteo Eq.(2.3), [11].

(3) L'ipotesi che  $h = \frac{1}{2J}\vec{A} \cdot \vec{A}$  e  $f(\vec{\alpha})$  sia un polinomio trigonometrico può essere sostituita da  $f(\vec{A}, \vec{\alpha})$  analitica e dipendente anche da  $\vec{A}$  e  $h$  a *Hessiano non singolare*:  $\det \vec{\partial}_{\vec{A}\vec{A}}^2 h(\vec{A}) \neq 0$ . Questo è il classico teorema KAM: può essere dimostrato con il metodo grafico qui esposto, [14], che differisce dal metodo originale di [3, 8] e dalla sua versione classica [4, 7].

(4) Il metodo qui esposto è tratto da [11] e fu ispirato da [9]: la differenza principale sta nella discussione dettagliata delle cancellazioni che in [9] sono invece fatte seguire da un'osservazione, si veda anche [14], che evita di esaminare i dettagli della serie di Lindstedt e deduce *a priori* che è nulla la somma dei contributi di tutti i termini che in [11] si suddividono in classi  $\mathcal{F}(\vartheta)$  che, come qui mostrato, sono individualmente nulle, *c.f.r.* Appendice. Nel metodo qui esposto gioca ruolo essenziale l'interpretazione della stima di Siegel in [10]. L'analogia con la teoria dei campi quantizzati e della rinormalizzazione é piuttosto chiara: si veda anche la versione in [15, Ch.8] in cui l'analogia è sfruttata in modo ancor più sostanziale e conduce a semplificazioni concettuali, [16].

(5) L'analiticità di  $h$  e  $f$  può essere sostituita da ipotesi di differenziabilità sufficientemente elevata (di ordine proporzionale a  $\ell$ , [5]). Una versione della dimostrazione di questo risultato può essere ottenuta con i metodi grafici qui esposti ma a prezzo di sostanziali ipotesi ulteriori, [17].

(6) Il metodo conduce anche a altri risultati nella teoria delle *risonanze*: ossia dei moti a rotazione  $\vec{\omega}_0$  per cui  $\vec{\omega}_0 \cdot \vec{\nu} = 0$  per uno o più vettori  $\vec{\nu} \neq \vec{0}$ . Abbastanza in generale si ottiene la esistenza di moti con queste frequenze di rotazione: ma si svolgono su tori invarianti di dimensione  $< \ell$  le cui equazioni parametriche  $\vec{\Delta}, \vec{\Xi}$  non sono analitiche a  $\varepsilon = 0$ : la serie di Lindstedt esiste



per questi moti ma è solo asintotica. Ciononostante con il metodo grafico si ottengono proprietà di analiticità di queste funzioni in domini complessi che hanno  $\varepsilon = 0$  sul bordo e che sono talvolta rappresentabili come serie di Borel o come serie frazionarie, [18, 19, 20]. Per una rassegna si veda [16].

(7) Oltre alla teoria delle risonanze il metodo è stato applicato allo studio di soluzioni periodiche e quasi periodiche di equazioni differenziali alle derivate parziali, [21, 22].

(8) Tra le applicazioni più rilevanti si annoverano le applicazioni alla Meccanica Celeste a cominciare da [23, 24].

(9) Infine esistono altre versioni della dimostrazione della convergenza della serie perturbativa per le equazioni parametriche dei moti quasi periodici: la proposta in [12] conduce a questa stessa conclusione attraverso un diverso algoritmo per la costruzione dei coefficienti che offre notevoli vantaggi dal punto di vista computazionale.

## A Appendici

(1) *Proprietà Diofantea*: Dato  $C$  e  $\tau > \ell - 1$  il volume dei punti  $\vec{\omega} \in R^\ell$  con  $|\vec{\omega}| < r$  che *non* soddisfano la Eq.(2.3) è stimato dal volume dello strato  $|\vec{\omega}_0 \cdot \vec{v}| < C^{-1}|\vec{v}|^{-\tau}$ ,  $\vec{v} \neq \vec{0}$  il cui volume si stima  $BC^{-1}r^{\ell-1}|\vec{v}|^{-\tau-1}$ , con  $B$  una costante opportuna (dipendente solo da  $\ell$ ). La unione di tutti questi strati ha volume  $\leq BC^{-1} \sum_{\vec{v} \neq \vec{0}} |\vec{v}|^{-\tau+1} = B'_\tau C^{-1}$  se  $\tau > \ell - 1$ . Quindi fuori da un insieme di volume 0 vale la proprietà Diofantea con una  $0 < C < \infty$ .

(2) *Teorema di Poincaré* (Versione semplificata):

Si supponga che per tutte le funzioni  $f(\vec{A}, \vec{\alpha})$  analitiche sia possibile trovare  $\vec{\Delta}_\varepsilon, \vec{\Xi}_\varepsilon$  in Eq.(1.5) analitiche in  $\varepsilon$  per  $\varepsilon$  abbastanza piccolo (in funzione di  $f$ ) e in  $\vec{A}', \vec{\alpha}'$  per  $\vec{A}'$  vicino a  $\vec{A}_0$ : allora la funzione generatrice  $\Phi_\varepsilon(\vec{A}', \vec{\alpha}')$  deve essere sviluppabile in serie di potenze di  $\varepsilon$ . Se  $\Phi^{(1)}(\vec{A}', \vec{\alpha}')$  è il contributo a  $\Phi_\varepsilon$  di ordine 1 in  $\varepsilon$ , questo deve verificare la  $\vec{\omega}(\vec{A}') \cdot \vec{\partial}_{\vec{\alpha}'} \Phi^{(1)}(\vec{A}', \vec{\alpha}') = 0$ . Ma in qualunque intorno di  $\vec{A}_0$  esistono  $\vec{A}'$  “risonanti”, ossia per i quali  $\vec{\omega}_0(\vec{A}') \cdot \vec{v} = 0$  per qualche  $\vec{v} \neq \vec{0}$ . Allora se  $f_{0,\vec{v}}$  ha  $\vec{v}$  fra le sue armoniche non nulle la  $\vec{\omega}(\vec{A}') \cdot \vec{\partial}_{\vec{\alpha}'} \Phi^{(1)}(\vec{A}', \vec{\alpha}') = f(\vec{A}', \vec{\alpha}')$  perché integrando membro a membro su  $\vec{\alpha}'$  dopo moltiplicazione per  $e^{-i\vec{\omega}(\vec{A}') \cdot \vec{v}}$  si troverebbe  $0 = \vec{\omega}(\vec{A}') \cdot \vec{v} \Phi_{\vec{v}}^{(1)}(\vec{A}') = f_{\vec{v}}(\vec{A}') \neq 0$  per  $\vec{A}' = \vec{A}_0$  (per esempio). Se si considerano le funzioni analitiche su  $U$  come elementi di uno spazio metrico con metrica  $\|f\| = \max_{|\Im \alpha_j| < \delta} |f|$  con  $\delta > 0$  allora le funzioni con *tutti* i coefficienti di Fourier non nulli sono

ovunque dense e quindi per esse non è possibile che il sistema di Hamiltoniana  $h(\vec{A}) + \varepsilon f(\vec{\alpha})$  sia integrabile analiticamente nel senso della Sez.1. Questo caso particolare si estende sulla base della stessa idea a casi più complessi, [10].

(3) *Deduzione della Eq.(3.7):*

Si procede constatando la validità delle seguenti identità che consentono di porre in forma più chiara i vari termini della Eq.(3.6):

(i) la prima parentesi quadra è di  $O(\varepsilon^2)$  mentre

(ii) nella seconda  $\varepsilon f(\vec{A}' + \varepsilon \vec{\partial}_{\vec{\alpha}} \Phi(\vec{A}', \vec{\alpha}), \vec{\alpha}) = \varepsilon f(\vec{A}', \vec{\alpha}) + O(\varepsilon^2)$  che si può scrivere  $\varepsilon(f(\vec{A}', \vec{\alpha}) - f_{\vec{0}}(\vec{A}')) + \varepsilon f_{\vec{0}}(\vec{A}') + O(\varepsilon^2)$ , come nella terza riga,

(iii) e la definizione di  $\Phi$ , che implica l'identità

$$\vec{\omega}_0 \cdot \vec{\partial}_{\vec{\alpha}} \Phi(\vec{A}', \vec{\alpha}) = -(f(\vec{A}', \vec{\alpha}) - f_{\vec{0}}(\vec{A}')) + (\vec{\omega}(\vec{A}') - \vec{\omega}_0) \cdot \sum_{\vec{v} \neq \vec{0}} \frac{i\vec{v} f_{\vec{v}}(\vec{A}')}{i(\vec{\omega} \cdot \vec{v})} e^{i\vec{v} \cdot \vec{\alpha}} \quad (\text{A.1})$$

e anche l'identità

$$\begin{aligned} (\vec{\omega}(\vec{A}') - \vec{\omega}_0) \cdot \vec{\partial}_{\vec{\alpha}} \Phi(\vec{A}', \vec{\alpha}) &= -(\vec{\omega}(\vec{A}') - \vec{\omega}_0) \cdot \sum_{\vec{v} \neq \vec{0}} \frac{i\vec{v} f_{\vec{v}}(\vec{A}')}{i(\vec{\omega} \cdot \vec{v})} e^{i\vec{v} \cdot \vec{\alpha}} \\ &+ \sum_{\vec{v} \neq \vec{0}} \frac{((\vec{\omega}(\vec{A}') - \vec{\omega}_0) \cdot \vec{v})^2 f_{\vec{v}}(\vec{A}')}{(\vec{\omega}_0 \cdot \vec{v})^2} e^{i\vec{v} \cdot \vec{\alpha}}, \end{aligned} \quad (\text{A.2})$$

il che consente di riscrivere il primo addendo nella seconda parentesi quadra, sommando le quantità in (ii),(iii), cosicché della seconda parentesi quadra rimane solo

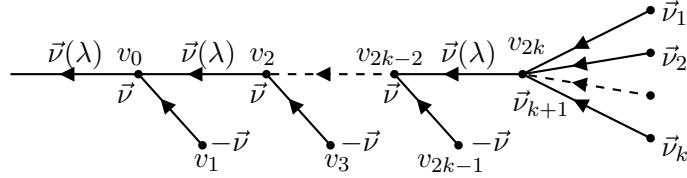
$$\varepsilon f_{\vec{0}}(\vec{A}') + \varepsilon (\vec{\omega}(\vec{A}') - \vec{\omega}_0) \cdot M_2(\vec{A}', \vec{\alpha}') (\vec{\omega}(\vec{A}') - \vec{\omega}_0) + O(\varepsilon^2) \quad (\text{A.3})$$

con  $M_2(\vec{A}', \vec{\alpha}')$  è una matrice  $\ell \times \ell$  analitica in  $\vec{A}'$  la cui definizione si legge dalle precedenti identità nelle quali  $\vec{\alpha}$  va espresso in  $(\vec{A}', \vec{\alpha}')$  via Eq.(3.3),

(iv) Infine  $\vec{\omega}_0$  in Eq.(A.3) può essere sostituito da  $\omega(\vec{A}'_0)$  con errore  $O(\varepsilon^2)$ , perché  $\vec{A}'_0 - \vec{A}_0 = \vec{B} = O(\varepsilon)$  ottenendo così la Eq.(3.7).

(4) *Serie di Lindsted e fattoriali:*

Si può avere  $\text{Val}(\vartheta) \simeq C^k k^{\alpha k}$  per un singolo albero  $\vartheta$  con  $3k + 1$  rami e opportune costanti  $C, \alpha$ . Ad esempio



se il momento  $\vec{v}(\lambda)$  che scorre nel ramo  $v_{2k-2}v_{2k}$  è di  $O(Nk)$ , e se i momenti estremi  $\vec{v}_j$  “cospirano” per produrre  $\vec{v}(\lambda)$  che saturo la condizione Diofantea  $C\vec{\omega}_0 \cdot \vec{v}_\lambda \sim (Nk)^{-\tau}$ . Allora si scelga  $\vec{v}_{2p} + \vec{v}_{2p+1} = 0$  per  $p = 0, \dots, k-1$  (così producendo  $k-1$  bolle singolari di scala  $k$ ) e si calcoli esattamente il valore di questo albero.

(5) *Proprietà Diofantea forte:*

Allo stesso modo il volume dei punti  $\vec{\omega} \in R^\ell$  con  $|\vec{\omega}| < r$  e che *non* verificano la proprietà Diofantea forte, Eq.(10.2), è stimato dalla somma di  $(|n|2^{-(n+4)^\ell})$  volumi di strati di spessore  $2^{n+1}|\vec{v}|^{-1}$  attorno ai  $2^{-(n+4)^\ell}$  punti  $\vec{v}$ : tale volume è  $\leq BC^{-1}|n|2^{n+1}r^{\ell-1}2^{-\frac{\ell}{\tau}}$  con  $B$  opportuna costante (dipendente da  $\ell$ ). Quindi il volume dei punti  $\vec{\omega}$  nella sfera di raggio  $r$  che *non verificano* la Eq.(2.3) o la Eq.(10.2) è stimato da una costante per  $\frac{1}{C}$  se  $\tau > \ell$  e quindi i punti che verificano sia la Eq.(2.3) che la Eq.(10.2) ha complemento di volume 0.

**Nota:** Sono grato ai colleghi che mi hanno dato suggerimenti per la revisione del manoscritto: in particolare Guido Gentile e Livia Corsi.

## References

- [1] T. Levi-Civita and U. Amaldi. *Lezioni di Meccanica Razionale*. Zanichelli, Bologna, 1927, ristampa 1974.
- [2] K. Siegel. Iterations of analytic functions. *Annals of Mathematics*, 43:607–612, 1943.
- [3] A. N. Kolmogorov. On the preservation of conditionally periodic motions. *In Lecture Notes in Physics, Stochastic behavior in classical and quantum Hamiltonians*, ed. G. Casati, J. Ford, Vol. 93, 1979, 93, 1979.
- [4] V. Arnold. Proof of A.N. Kolmogorov’s theorem on the preservation of quasi-periodic motions under small perturbations. *Russian Mathematical Surveys*, 18:13–40, 1963.

- [5] J. Moser. On invariant curves of an area preserving mapping of the annulus. *Nachrichten Akademie Wissesbhaftes Göttingen*, 11:1–20, 1962.
- [6] H. Poincaré. *Les Méthodes nouvelles de la mécanique Céleste*. Les grands classiques Gauthiers-Villars. J.Gabay, Paris, 1987.
- [7] G. Gallavotti. Perturbation theory for classical Hamiltonian systems. *in Scaling and self similarity in Physics*, Ed. ed. J. Fröhlich, Birkhauser, Boston:359–426, 1985.
- [8] G. Benettin, L. Galgani, A. Giorgilli, and J. Strelcyn. A proof of Kolmogorov's theorem on invariant tori using canonical transformations defined by the Lie method. *Nuovo Cimento B*, 79:201–223, 1984.
- [9] L.H. Eliasson. Absolutely convergent series expansions for quasi periodic motions. *MPEJ (Mathematical Physics Electronic Journal)*, 2, n.4:1–33, 1986.
- [10] J. Pöschel. Invariant manifolds of complex analytic mappings. *Phénomènes Critiques, Systèmes aleatoires, Théories de jauge, Proceedings, Les Houches, XLIII (1984), North Holland, Amsterdam, II:949–964*, 1986.
- [11] G. Gallavotti. Twistless KAM tori. *Communications in Mathematical Physics*, 164:145–156, 1994.
- [12] A. Giorgilli and U. Locatelli. Kolmogorov theorem and classical perturbation theory. *ZAMP: Zeitschrift für angewandte Mathematik und Physik*, 48:220–261, 1997.
- [13] G. Gallavotti and G. Gentile. Majorant series convergence for twistless kam tori. *Ergodic Theory and Dynamical Systems*, 15:857–869, 1995.
- [14] G. Gentile and V. Mastropietro. Methods of analysis of the Lindstedt series for KAM tori and renormalizability in classical mechanics. A review with some applications. *Reviews in Mathematical Physics*, 8:393–444, 1996.
- [15] G. Gallavotti, F. Bonetto, and G. Gentile. *Aspects of the ergodic, qualitative and statistical theory of motion*. Springer Verlag, Berlin, 2004.
- [16] G. Gentile. Quasi-periodic motions in dynamical systems. Review of a renormalisation group approach. *Journal of Mathematical Physics*, 51:015207 (+34), 2010.
- [17] F. Bonetto, G. Gallavotti, G. Gentile, and V. Mastropietro. Lindstedt series, ultraviolet divergences and Moser's theorem. *Annali della Scuola Normale Superiore di Pisa*, 26:545–593, 1998.

- [18] O. Costin, G. Gallavotti, G. Giuliani, and G. Gentile. Borel summability and Lindstedt series. *Communications in Mathematical Physics*, 269:175–193, 2006.
- [19] G. Gallavotti and G. Gentile. Degenerate elliptic resonances. *Communications in Mathematical Physics*, 257:319–362, 2005.
- [20] G. Gallavotti, G. Gentile, and A. Giuliani. Fractional Lindstedt series. *Journal of Mathematical Physics*, 47:012702 (+33), 2006.
- [21] G. Gentile and M. Procesi. Periodic solutions for a class of nonlinear partial differential equations in higher dimension. *Communications in Mathematical Physics*, 289:863–906, 2009.
- [22] G. Gentile and V. Mastropietro. Construction of periodic solutions of the nonlinear wave equation under strong irrationality conditions by the Lindstedt series method. *Journal de Mathématiques Pures et Appliquées*, 83:1019–1065, 2004.
- [23] V. Arnold. Small denominators and problems of stability of motion in classical and celestial mechanics. *Russian Mathematical Surveys*, 18:85–191, 1963.
- [24] V.I. Arnold and V.V. Kozlov and A.I. Neishtadt. *Mathematical aspects of classical and celestial mechanics*, volume 3 of *Encyclopaedia of Mathematical Sciences*. Springer, Berlin, 2002.